Bruce Aune

An Empiricist Theory of Knowledge



Copyright ©2008 Revised 3/11/09

Bruce Aune

To Anne

CONTENTS

PREFACE iv

Chapter One: WHAT IS KNOWLEDGE?

Conceptions of Knowing 1
Epistemic Contextualism 4
Lewis's Contextualism 6
A Dual Analysis of Knowledge 11
Problems for Two Senses of "Knows" 14
Avoiding Gettier Counterexamples 15
Concluding Remarks 19

Chapter Two: THE CLAIMS OF RATIONALISM 23

The A Priori, Universality, and Necessity 23
Axioms and Primitive Rules of Inference 26
General Doubts about Intuitive Knowledge 28
Logical Truths and Rules of Inference 32
Alleged Self-evident Factual Truths 36
Three Final Examples, Two Old and One New 40
An Indirect Argument for Rationalism 43

Chapter Three: EMPIRICISM AND THE A PRIORI 47

Quine's Criticism of Analyticity 47
Quine's Later View of Analytic Truth 50
Analyticity, Logic, and Everyday Language 52
Analyticity Extended 57
Some Examples and Arguments by Kripke 63
Beliefs, Propositions, and Analyticity 64

Chapter Four: PROPERTIES AND CONCEPTS 71

What are Properties? 71
Problems with A-theories and T-theories 73
Predication 76
Advantages of F-theories 78
What are Concepts? 80
Some Problems about DSTs 83
More about Concepts 85
Concepts, Predicates, and the World 87
Meaning, Intending, and Content Clauses 88
Concluding Remarks 93

Chapter Five: OBSERVATIONAL KNOWLEDGE 95

A Problem about Observation 95
Evaluating Observational Beliefs 97
Does Knowledge Need a Foundation? 99
Alternatives to Foundationalism 102
Knowledge and the World: Some Problems 104
Semantic Externalism 107
Criticism of Semantic Externalism 110
A Skeptical Problem Restated 114

Chapter Six: MEMORY AND A POSTERIORI KNOWLEDGE 115

Memory as a Source of Knowledge 115
What is Induction? 122
Induction: Arguments Pro and Con 124
Induction and Laws 126
Inference to the Best Explanation 128
Inferences Based on Bayes' Theorem 130
Ascertaining Prior Probabilities 133
Basic Prior Probabilities 136
The BIV Hypothesis Again 138
Concluding Remarks 142

APPENDICES 145

1. Armstrong's New Hypothesis About Universals 145
2. Boghossian and Field on Basic Logical Principles 146
3. Stipulation, Proper Definitions, and Truth 149
4. "What is Said" and Propositions 152
5. Chisholm's Defense of the Principle CP 155
6. Analytic Probability Principles 156

REFERENCES 159

INDEX 162

Preface

When I began to teach philosophy, almost every responsible analytic philosopher was an acknowledged empiricist. Today, many analytic philosophers repudiate central tenets of the position, rejecting at least an analytic/synthetic distinction and often pursuing metaphysical questions that main-line empiricists set aside decades ago. I regard this development as unfortunate, a backward step in philosophy that needs to be corrected. Bas van Fraassen and Anil Gupta have recently taken important steps in the right direction. I try to do my part in this book, attacking well-known criticisms of empiricist doctrine and defending the sort of empiricist theory that I consider acceptable.

There is no essence to empiricism: different positions have been defended under the name "empiricism," and the practice will no doubt continue.[1] The empiricism I learned as a student was called "logical empiricism," the qualifier marking the importance of formal logic to this version of the theory. One of my teachers, Herbert Feigl, discussed the distinctive claims of an earlier form of this empiricism in a programmatic article called "Logical Empiricism," which was regarded as something of a manifesto in its day. The adjective Feigl attached to "empiricism" was chiefly owing to the logical and semantical work of Rudolf Carnap, another of my teachers,[2] who was the dominant figure among the empiricists whose views Feigl was promoting. Carnap's epistemological views changed significantly over his long philosophical career, and his later views represent a development of empiricism that deserves to be better known by today's critics of the doctrine.

The objections to empiricism that W. V. Quine formulated in "Two Dogmas of Empiricism" (1951) are still widely regarded as successful. The supposed dogmas in question are, Quine said, a belief in a fundamental cleavage between analytic and synthetic truths, on the one hand, and reductionism, the belief that each meaningful statement is equivalent to some logical construct upon terms referring to immediate experience, on the other. As it happens, logical empiricists did not hold the second dogma when Quine's paper was published. Feigl explicitly denounced it in "Logical Empiricism" (first published in 1943) and Carnap left it behind in the middle thirties.[3] The first dogma—"assumption" is really a better word here—was indeed accepted, at least as an ideal, by logical empiricists, and I shall therefore discuss Quine's objections to it carefully and at length. Although Quine has long been one of my philosophical heroes, I have to say that his objections to analyticity fail to undermine the position Carnap defended in his later years. Carnap was right to set them aside as unsuccessful.

Quine's criticism was not the only cause of empiricism's decline; another was the rise of epistemological rationalism. The cause of this phenomenon is complex; no doubt it had something to do with the revival of interest in metaphysics that resulted from Kripke's revolutionary ideas on identity, necessity, and essential properties. But whatever the actual cause may be, the most influential exponent of the new analytical rationalism turned out to be R. M. Chisholm. As Alvin Plantinga remarked in 1990, Chisholm's thought "has [in fact]…dominated American epistemology for more than thirty years";[4] if this is an exaggeration, as I believe it is, it is nevertheless not very far from the truth. To defend a version of empiricism at the present time it is therefore not sufficient to overcome the criticism of Quine;

---

[1] See van Fraassen (2002), Appendix B, "A History of the Name 'Empiricism'."
[2] As a graduate student, I spent a year at UCLA, where he was then teaching.
[3] He rejected it in Carnap (1936).
[4] Plantinga (1990), p. 366.

one should also criticize the arguments supporting the alternative position that Chisholm was instrumental in initiating.

Criticizing an alternative position is unfortunately an awkward task. There is always more than one version of such a position and always more than one advocate to confront. Alvin Plantinga, George Bealer, Laurence BonJour, Christopher Peacock, and Robert Audi have defended well-considered versions of epistemological rationalism, but I cannot examine all of them in a book like this. Instead of focusing attention on particular versions of the doctrine, I shall for the most part attend to what I regard as the most important arguments rationalists offer for synthetic a priori truths. These arguments feature a number of examples that are cited again and again; in 2005 Laurence BonJour offered the examples that Chisholm gave as early as 1966.[1] I therefore attend to them closely. My positive arguments against epistemological rationalism depend not on the peculiarities of different rationalist theories but on structural weaknesses common to them all. The propositions they take to be intuitively decidable synthetic truths are actually warranted, if they are actually true, by facts that are far too discursive and stipulative to be a confirming element in any rationalist theory.

In the past twenty years or so academic philosophy has become highly specialized, with the result that philosophers working in epistemology often do not have well-considered views in related subjects such as metaphysics, formal semantics, and philosophy of science. Topics in these related subjects are nevertheless crucially important for basic epistemological disputes. As far as a priori knowledge is concerned, the pertinent topics belong mainly to metaphysics and formal semantics. Specific issues concern the reality and nature of properties and propositions, which rationalists typically regard as providing the foundation for synthetic a priori knowledge. Because recent work in metaphysics and formal semantics puts older views of these supposed objects into serious doubt, I devote part of one chapter to propositions and a whole chapter to properties. Writing this material has reinforced my belief that it is absolutely essential for a responsible treatment of a priori knowledge.

Although critics of empiricism have typically concentrated on an analytic/synthetic distinction, a satisfactory empiricist philosophy must provide an acceptable account of a posteriori knowledge. In my final chapters I therefore discuss problems with the sources of empirical knowledge that empiricists almost always accept: observation, memory, and what Hume called experimental inference. I open chapter five with a consideration of some of these problems, but I soon address the doctrine of semantic externalism that Hilary Putnam developed in criticizing his well-known "brains in a vat" version of a perennial skeptical hypothesis. Although Putnam evidently considered his externalism to be opposed to traditional views of meaning and reference, I argue that it is in fact quite close to the verificationism that was espoused by logical positivists, and I reject it for reasons that apply to that once popular doctrine. As I see it, the empiricists' historical repudiation of empirical entities that cannot possibly be observed is something that an acceptable empiricism must leave behind.

In chapter six I am mainly concerned with "inductive" or a posteriori inference, which has been seriously neglected by main-line epistemologists with rationalist sympathies. (Chisholm had almost nothing to say about this kind of inference in the last edition of his influential *Theory of Knowledge*.) Since inductive methods raise more problems than most philosophers seem to realize, I provide a critical overview of the standard alternatives. My assessment of these methods is generally negative—even for the current favorite, Inference to the Best Explanation. Arguing

[1]See Chisholm (1989) and BonJour (2005). Chisholm used the same examples in earlier editions of his *Theory of Knowledge*; the first was published in 1966.

that the problems familiar methods are supposed to solve can be disposed of only by relying on Bayes' theorem of probability theory, I end up discussing this theorem and its relation to what can be called "evidential" probability. This kind of probability is often viewed as a measure of subjective belief, but I argue that it must be understood differently if well-known problems are to be avoided. I view it as a measure of certainty and evidential support, a position I do my best to justify. Although some epistemologists are very knowledgeable about probability theory, the subject is evidently daunting to many philosophers. Believing this, I took special pains to make my discussion understandable to those entirely new to the subject. Sophisticated readers can simply skip the explanatory passages I include here and there.

To deal with certain side issues that are highly significant for some philosophers but of minimal interest to many, I followed the example of Fogelin (1994) and van Fraassen (2002) and included a number of appendices. These appendices are generally too long to be footnotes but they are well suited to the end of the book where readers who recognize their importance can consult them. Each is concerned with matters that, in my experience, always eventually arise when philosophical rivals debate epistemological issues.

A number of friends contributed to the manuscript in one way or another. Joe LaPorte, Steve Braude, Jeffrey Sicha, and Lynne Baker made helpful comments on the chapters they read. LaPorte was particularly helpful with chapters one and six, and Sicha subjected the whole manuscript to very careful scrutiny, doing the sort of thing he did many years ago when I was writing my first book. I am greatly indebted to his good judgment and critical acuity. The late Gregory Fitch offered illuminating remarks on some questions I had with Kripke's footnote about the "necessity of origins"; he was not himself critical of the argument Kripke seemed to give, but his remarks were instrumental in leading me to the criticism I formulate in chapter three. My wonderful wife, Anne, to whom I dedicate this book, was helpful from start to finish. She read every version of the manuscript and always discovered errors that I had somehow missed.

This is my fourth book on epistemology. The first was principally indebted to the work of Wilfrid Sellars, whose influence is discernible here mainly in chapters four and five. The logical empiricist doctrines that I absorbed from Herbert Feigl, Rudolf Carnap, and Grover Maxwell are evident in chapters two and three, and the logical and semantic theory I learned from Donald Kalish and Richard Montague is also apparent there. Montague introduced me to the logical foundations of probability, but my views on that subject are more strongly indebted to the writing of my one-time students Roger Rosenkrantz and Brian Skyrms. The fact that these people, and certain writers whom I have not mentioned, do not agree on all philosophical matters may help to explain the independence of my own philosophical thinking, such as it is. I have had no single path to follow.

Chapter 1

WHAT IS KNOWLEDGE?

Most of the topics I shall be concerned with in this book concern kinds of knowledge: *a priori* knowledge, observational knowledge, and the possibility of having knowledge about objects and processes that, like the feelings of others or the micro-objects of current physics, cannot possibly be perceived. But these kinds of knowledge make sense only as instances of knowledge in a generic sense. Plato discussed knowledge of this generic sort in his *Theaetetus*, the first analytical study in epistemology, and the nature of knowledge so understood has been the most widely discussed topic in recent work on the subject. Since current disputes about the general nature of knowledge are closely tied to competing strategies for making progress in philosophy, a discussion of this topic is an appropriate starting point for the argument of this book. What I say here provides a foundation for what I shall argue in later chapters.

Conceptions of Knowing

An analytical study of knowledge ought to acknowledge that the word "knowledge" is significantly ambiguous—as are its equivalents in other languages, such as the Greek epistêmê, from which "epistemology" is derived. The principal meanings of these words can be arranged into three groups. The first group concerns abilities of various kinds, primarily cognitive abilities that result from learning but sometimes even motor abilities. One can know German or know how to walk on stilts; one can know how to give a rousing speech, how to use the library, how get to the airport, but also how to do a handstand or back flip. Another group involves acquaintance, familiarity, personal experience, and corresponding recognitional abilities. One can know a former teacher; one can know a person by name or by sight; one can know fear, love, or disappointment; and can know New York, Boston, or the neighboring university campus. The last group of meanings—perhaps it is a single meaning—concerns "facts gathered by study, observation, or experience," and conclusions inferred from such facts (as when one has an in-depth knowledge of particle physics)."[1] What the dictionary describes as knowledge of facts can be described more plainly as knowledge-that:[2] knowledge that snow is while, that grass is green, or that 2+2 = 4. It is this last sort of knowledge that is central to recent work in epistemology.

In the early part of the last century some philosophers, notably Bertrand Russell, considered acquaintance or direct experience the fundamental source of empirical knowledge; for them, knowledge-that ultimately arises from knowledge of. As they saw it, our subjective experiences are elements of our consciousness, and everything we know by perception arises from our experiences. This view is no longer widely held: most philosophers now contend that acquaintance involves a substantial amount of knowledge-that, and the directly experienced residue in experience is little more than a stimulus for interpretive acts that result in more knowledge-that. Just think of your knowledge of your own hometown. You know that it has various buildings, various streets, various parks; you know where your house or apartment was—you know that it was in such and such a place. You can call up memory images of

---

[1] See the entry under "knowledge in" *The Oxford American Desk Dictionary and Thesaurus.*

[2] I here assume that knowledge-who, knowledge-what, knowledge-when, and so on, are special cases of knowledge-that. For example, "Tom knows who wrote the *Declaration of Independence"* attributes to Tom the knowledge that X wrote the *Declaration of Independence,* where X is the person, namely Thomas Jefferson, who wrote that document.

places you recall, but these images simply bring more facts to mind. The prevalence of this new view of acquaintance—the idea that it is not a distinctive kind of knowledge more basic than knowledge-that—owes a lot to Wittgenstein's attack on what he called "private languages," [3] and it may or may not be right or defensible. I shall have more to say about acquaintance in chapter five.

Before 1963 analytically-minded philosophers mostly agreed that knowledge-that could be understood as justified true belief. Edmund Gettier's now famous criticism of this account destroyed the agreement and stimulated a plethora of attempts to provide an improved definition.[4] The philosophers seeking an improvement had two desiderata specifically in mind. They wanted a definition incorporating standards that would make it possible for ordinary human beings to know most of what they think they know, and they wanted a definition that would avoid Gettier examples and others relevantly like them. A definition having the first feature would be instrumental in avoiding skepticism, an outcome that could be expected if the required standards of evidence were set too high. They also assumed that a definition having the desired features would require a knower to possess an appropriate true belief.

The great number and variety of attempts to provide a definition satisfying the desiderata I mentioned[5] make it fairly clear that the philosophers attempting to provide such an improvement were not working with a single knowledge concept that already existed and was generally accepted. They may have had illusions about what they were doing, but the reality is that they were attempting to *create* a knowledge concept that was philosophically preferable to the simple one that Gettier criticized. They wanted a better analytical account of what knowledge could be taken to be. As it happened, they did not definitely succeed in this endeavor: no generally accepted conception or account of the desired kind was ever created. Many philosophers continue with the hunt, but some have basically given up on it. Among the latter, Timothy Williamson came to the conclusion that "knowing does not factorize as standard analyses require."[6] Instead of attempting to provide a definition of knowledge, Williamson offered a "modest nonreductive analysis," describing knowing as "the most general factive, stative [human] attitude"—factive in being attached only to truths, and stative in being a state rather than a process.[7] But Williamson's nonreductive analysis does not appear to have attracted many adherents. Most philosophers appear to want a more informative account of knowing than Williamson's analysis provides.[8]

The consensus that once existed on seeking an improved justified-true-belief (or JTB+) analysis of knowing broke down for other reasons. Some philosophers, such as Peter Unger and Robert Fogelin, did not believe that skepticism should be ruled out by easily satisfied standards for knowing. These philosophers even wrote books supporting versions of that generally abhorred doctrine.[9] In taking a skeptical

---

[3] See Wittgenstein (1959),§§ 243- 272. I critically assess Wittgenstein's argument in Aune (1991), pp. 115-130.

[4] See Gettier (1963).

[5] The principal definitions given in the first twenty years after Getter's paper was published are ably discussed in Shope (1983).

[6] See Williamson (2000), p. 4.

[7] Ibid, p. 34.

[8] Hilary Kornblith (1999) defends a singular account of knowing that I do not consider here. He argues that knowledge is a natural kind, instances of which may be possessed by birds or monkeys as well as human beings. I do not deny that a knowing concept with a wide application of this kind is possible, but like Hacking (2005), I find problems in the very concept of a natural kind, and I think the word "knowing" is in any case applied to a more diverse variety of instances than is happily accommodated by Kornblith's single conception. As I see it, when we describe birds or monkeys as knowing things, we are using the word "knowing" in an extended, analogical sense.

[9] See Unger (1975) and Fogelin (1994) and (2000).

line they had little trouble satisfying the other desideratum for a JTB+ analysis of knowledge, the one requiring the avoidance of Gettier examples. Each of the examples Gettier actually gave presupposed that a person may know that P on the basis of inconclusive evidence--evidence that does not exclude the possibility that P is actually false.[10] But supporters of skepticism normally endorse higher standards for knowing: they seek evidence that is logically conclusive.[11] Since a skeptical scenario featuring Descartes' evil genius or Putnam's brains in a vat cannot be conclusively refuted (or ruled out with utter certainty) by any evidence plausibly available to an observer, a philosopher requiring conclusive evidence for knowing will end up with the view that no alternative scenario incompatible with skepticism can possibly be known to be true.

Thus far I have been speaking of assumptions about knowledge that philosophers have held since 1963. Before that further differences existed, particularly if we go back far enough. Plato held that knowledge (epistêmê) is infallible and, unlike belief, directed to an immutable object.[12] Aristotle held knowledge to be either immediately certain or a demonstrative consequence, via the syllogism, of immediately certain premises.[13] Descartes did not limit necessary inference to the syllogism, but like Aristotle he thought properly scientific knowledge, or *scientia,* required rational certainty: the subject's evidential basis for such knowledge must be conclusive.[14] Earlier twentieth-century philosophers had a more flexible attitude to knowing. G. E. Moore held that "I know that P" sometimes does, and sometimes does not, imply "I know that P with utter certainty";[15] and in 1952 Norman Malcolm distinguished a strong from a weak sense of "knows," one implying that the subject is certain of something, the other not.[16]

In everyday life we often apparently do speak of knowledge in what Malcolm called the weak sense; we seem to assume that people often have genuine knowledge when their evidence is logically inconclusive, when it does not exclude the possibility of error. We seem to assume this when, having looked at our watch, we say we know what time it is; we seem to assume it when, watching a television newscast, we say we know the Twin Towers have been destroyed by a terrorist attack; and so on. But sometimes we speak of it in what is pretty clearly a stronger sense, one requiring that a subject's evidence be logically conclusive or very close to it. (One way of describing logically conclusive evidence is to say that E is logically conclusive for P when the evidential probability of P on the basis of E is 1, an idea I ex-

---

[10]This is easily seen. In setting forth his counter-examples, Gettier described cases in which a subject, S, has an adequately justified belief that P, which nevertheless happens to be false. Not knowing that P is false but knowing elementary logic, S forms the belief that Q, and this belief, because of its known logical relation to the adequately justified P, is adequately justified as well. But Q, unlike P, happens to be true, and true for reasons having nothing to with the evidence S possesses. S therefore satisfies the justified-true-belief conditions for knowing that Q, but because of the logically fortuitous character of Q's truth, S clearly does not possess this knowledge.

[11]The adjective "logical" here is customary but it is not really happy, since P can provide conclusive evidence for Q without it being a *logical* truth that P only if Q. It is not a logical truth that if Sarah is a sister, Sarah is female, but the former provides conclusive evidence for the latter. I pursue matters of this kind in chapter three.

[12]*Republic* 511d.

[13]*Posterior Analytics* I, 71b.

[14] In *The Principles of Philosophy* Descartes described this as "perfect" knowledge; his Latin equivalent was "scientia"; see Cottingham, *et al* (1985), vol. 1, pp. 10n and 179. In addition to the "absolute certainty" provided by perfect knowledge, Descartes accepted a conception of moral certainty, which is close to what Malcolm (see note 14 below) considered knowledge in the weak sense; see Cottingham, ibid. p. 290.

[15]Moore (1959), p. 236f.

[16]Malcolm (1963). Malcolm's position is actually more complicated than I indicate in the text; it involves qualifications that are difficult to spell out in a brief statement.

plain fully in chapter six.) In a recent letter to the *Scientific American*, a man calculated that to win the $160-million with his lottery ticket, he would have to beat the winning odds of 1 to 120,526,770. In spite of these odds, he was willing to buy the ticket, and when he bought it we would not agree that if his friend Tom believes he will lose, Tom knows he will lose if that is what will happen. In spite of the very strong evidence Tom possesses, the possibility remains that the man will win--and this is enough to defeat Tom's claim to know he will lose. In this case, actually knowing that the man will lose seems to require rational certainty: our evidence must be sufficient to rule out the possibility that he will win.

The idea that we do in fact commonly apply different standards of evidence or different levels of certainty in deciding whether this or that person has knowledge under these or those circumstances is now widely accepted,[17] but some philosophers give "invariant" accounts of this diversity.[18] According to some, knowledge-ascriptions based on weak standards are usually in fact false, though they may have some practical value;[19] according to others, negative ascriptions ("S does not know that P") based on exceptionally strong standards are actually false, though they seem plausible in the context of some well-known skeptical arguments.[20] The key issue in the whole debate is how the diversity that is apparent in assertions involving "knows that" is best accommodated theoretically, and what account of how knowledge may be understood is most illuminating. As it happens, I shall be defending a dual account in what follows, one in which a concept of knowing for certain is distinguished from a minimal concept that does not require rational certainty. My approach is not widely accepted at the present time, however; the most widely discussed alternative in recent years is some form of contextualism. Because of its popularity as well as its complexity and suggestiveness, I want to consider this sort of view first.

Epistemic Contextualism

Although the term "contextualism" has been applied for more than a decade to the view that ascriptions of the form "s knows that p" are properly evaluated by stronger or weaker standards in differing contexts, some writers have recently emphasized that this view is more aptly described as "epistemic relativism."[21] The new terminology is supported by the consideration that many knowledge ascriptions whose truth-values differ in different contexts do so for reasons having nothing to do with varying epistemic standards.[22] A representative example is "Tom knows that George is six feet tall," which, since "knows that P" implies "P," conveys the idea that George is six feet tall at the time Tom is said to have this knowledge. But George's height changes over the course of his life. For most of his boyhood George is far from tall; at maturity he may be six feet tall; and as an old man he will be shorter than this. So if George is six feet tall when Tom is said to know he has that height, the knowledge ascription is true; if he is taller or shorter when Tom is said to be this tall, the knowledge ascription is false. As a general matter, the indexical elements—the pro-

[17] See, e.g., DeRose (1995), Lewis (1996) and Cohen (2000).
[18] This adjective is commonly used to identify the opponents of epistemic contextualists. See Conee (2005) or Bach (2005). Macfarlane also uses the adjective but he distinguishes two kinds of invariantism, strict and sensitive, only the former being incompatible with contextualism. See Macfarlane (2005), p. 199.
[19] Unger (1975), Fogelin (1994)
[20] Bach (2005).
[21] See Macfarlane (2005), offers an illuminating taxonomy of recent views about the semantics of "know."
[22] Feldman (1999) may have been the first to emphasize this.

nouns, tensed verbs, and other contextual indicators—in both the "that"-clause and the words preceding it in a knowledge ascription (for instance in "Tom once knew") may have a decisive effect on the ascription's truth, and this effect has nothing to do with varying standards of evaluation.

Are there clear cases in which different utterances of a knowledge-ascription type are rightly evaluated by epistemic standards of varying stringency? The word "rightly" is the crucial modifier here. Contextualists or epistemic relativists say yes; invariantists, as they are sometimes called, say no. How are we to decide who is right? Or is there perhaps no fact of the matter to be right about? The differing parties here obviously have access to the same linguistic or behavioral data. They might not, of course, attend to all the data equally well. I have shown that different philosophers have had in the past, and have now, different convictions about the nature of knowledge; and recent investigation shows that different groups of non-philosophers—and sometimes even the same ones—speak about knowledge in inconsistent ways.[23] Philosophers almost always say that knowing that P implies it is true that P, but ordinary people sometimes say that they have known things that turned out to be false.[24] Similarly, although most philosophers insist that knowing that P implies believing that P, David Lewis rejects this implication, building his conception of knowledge on its denial.[25] Contextualists and invariantists (whether skeptics or dogmatists) who argue about the plurality of proper or acceptable epistemic standards augment these instances of disagreement. A plethora of varying usages also turn up in Google searches focused on "'knows' and evidential standards" and "'knows' and certainty."[26] The only reasonable conclusion to draw from these incompatibilities in belief and usage, it seems to me, is that there is really no single objective fact of the matter—no single property, concept, or standard—that is available to prove that one position in the debate is right and the others wrong.

I hasten to add that even if there is no decisive fact of the matter here, one position may nevertheless be philosophically more satisfactory, all things considered, than the others. Respecting existing usage is not a decisive requirement for an acceptable philosophical analysis or conceptual clarification. Some usage is clearly more discerning than others; some is better-informed and more relevant to philosophical issues than others; and some is even inconsistent, raising more problems than it solves. As I shall argue at length in chapter three, philosophical analysis is inherently and inevitably revisionary or, to use Carnap's term, reconstructive.[27] If contextualism is preferable to the alternatives I have mentioned, it is so only because it can be spelled out in such a way that it succeeds in resolving pertinent philosophical issues more satisfactorily than those alternatives are capable of doing.

Two matters that should be explained by a satisfactory contextualist (or epistemically relativist) theory are (a) what, according to it, knowledge is or consists of and (b) how the alternative epistemic standards it postulates are to be identified. The only contextualist theory so far developed that deals with both matters in a detailed way is the one developed by David Lewis. I shall therefore comment briefly on the basic elements of his view. I shall, as I implied, reject his contextualism, but I shall nevertheless accept some of the key ideas on which it is based. Although the distinction I mentioned between the context-dependence owing to a formula's indexical features and its alleged susceptibility to evaluation by stronger or weaker

---

[23] See Nichols, Stitch, and Weinberg (2003).

[24] Bach (2005), p. 62. Jay Rosenberg is a philosopher who comes close to siding with the ordinary people on this point; he holds that "S knows that P" is consistent with "Not-p." See Rosenberg (2002), pp. 1-2. I discuss Rosenberg's view thoroughly in "Rosenberg on Knowing" (in preparation).

[25] Lewis (1996).

[26] See Ludlow (2005).

[27] See below, footnote 41.

AvE4eVaMuViMixRecords WELCOMME AvE4eVaMuViMixRecords

standards needs to be incorporated into Lewis's theory, I shall ignore it here. It is not pertinent to the issues that concern me.

Lewis's Contextualism

Instead of holding that there is more than one sense of "knows," Lewis says that the formula "S knows that P" can be given a single definition by means of which we may ascertain the truth-conditions for utterances conforming to it in this or that context. If "knows that P" is truly ascribed to a subject S in a context C, S must possess evidence, Lewis says, that eliminates every alternative possibility relevant in C. Possibilities relevant this way have two distinguishing features: they include ~P and they are properly ignored in C. Lewis does not intend that a subject's evidence should eliminate the possibilities including ~P at one fell swoop, by directly supporting the truth of P, which is incompatible with these possibilities. He intends that the evidence should eliminate each relevant not-P possibility directly; and as the result of eliminating all these possibilities, it will thereby support the proposition P as the only remaining alternative.

To understand Lewis's position fully, we need to know what he means by evidence, how he thinks evidence can rule out a possibility, and how he identifies the possibilities that are relevant in a given epistemic context. As for a subject's evidence, Lewis takes this to be the subject's "entire perceptual experience and memory."[28] If I were observing an Airedale terrier, I would have a characteristic perceptual experience, one different from what I would have if I were observing a tiger, a phone booth, or a Volkswagen beetle. Of course, my perceptual experience in observing an Airedale is not itself sufficient for knowing that what I see is an Airedale; I must have some background knowledge about Airedales and other things I might be observing. I possess this knowledge because I remember what I have previously learned about these things and what I have experienced in connection with them. What I perceive in a given context and everything I remember pertinent to it is the evidence at my disposal for the case at hand.

The evidence, thus understood, that I have in a given context rules out any relevant possibility in which my entire perceptual experience and memory are not as they are in actuality. Lewis understands actuality to be the possibility that actually obtains. My evidence thus rules out any relevant possibility that does not match actuality with respect to my entire perceptual experience and memory. If a possibility does match reality in this way, it is *uneliminated* by my evidence. It is important to emphasize that some possibilities that do match actuality with respect to my evidence may yet be ruled out because they are deemed *not relevant* to the subject at hand. My evidence may match the possibility that I am being deceived by Descartes' evil demon, but that possibility can normally be disregarded as not relevant to the question of what I now know: normally, it is not a possibility that must be ruled out by my evidence here and now.

Lewis provides seven rules for identifying possibilities that may or may not properly be ignored in a given context. Since the criticism I shall make of Lewis's definition does not depend on the details of these rules, I shall describe them briefly and not comment on subtleties pertinent to them. I shall say just enough to give the reader a sense of how they may be deployed in responding to objections that might be raised against Lewis's definition.

---

[28]Lewis (1999), p. 424.

Four of Lewis's rules identify possibilities than cannot properly be ignored. The first is the *Rule of Actuality*. According to this rule, a possibility that actually obtains is never properly ignored. A possibility that actually obtains does not, of course, have to be eliminated in showing that someone knows something, but in attending to it one may become obliged by other rules to consider further possibilities that might have to be eliminated. The other three rules have this effect. According to the *Rule of Belief*, we cannot ignore anything the subject believes to obtain or, given his available evidence, should believe to obtain. Since what a subject believes or should believe to obtain may occasionally be at odds with his or her actual evidence (something belonging to actuality), the Rule of Belief may identify possibilities that have to be eliminated in deciding what the subject actually knows. The *Rule of Resemblance* introduces further possibilities of this kind. If two possibilities saliently resemble one another, then if one may not properly be ignored because of a rule other than this one, the other may not properly be ignored either. Lewis qualifies[29] this rule by adding that the salience of the relevant similarity should pertain mainly to the subject's evidence: if a possibility differs radically from actuality except for its resemblance to the subject's evidence, the Rule of Resemblance does not apply. This qualification has the effect of opposing skepticism. If my total evidence when I am actually perceiving a brilliant sunrise saliently resembles the evidence I would have if I were being deceived by Descartes' evil demon, the radical difference between my actual situation and the deceptive one renders the Rule of Resemblance inapplicable in this case. Another rule could conceivably require me to eliminate this skeptical possibility, but the Rule of Resemblance would not require me to do so. The final rule, *Attention,* has the effect of making knowledge elusive in philosophical contexts. It says that a possibility that is not ignored is not properly ignored, no matter how likely it may be in view of the evidence.

Lewis's last three rules tell us what we may properly ignore in determining whether someone knows something. Like his qualification to the Rule of Resemblance, the first three of these rules provide impediments to unbridled skepticism. Lewis calls the first one the Rule of Reliability. According to this rule, perception, memory, and the testimony of others may be considered generally reliable; as a result, we may—"defeasibly," Lewis says—ignore possibilities in which they fail. (In saying that ignoring these possibilities is defeasible, Lewis means that the presumption that these sources of knowledge are reliable may be defeated, or overridden, by evidence that casts doubt on them in a particular case.) Lewis's second rule in this group concerns Permissible Rules of Method. According to it, we may assume, defeasibly, that our evidence samples are representative and that the "best explanation" of our evidence—the available explanatory account that, if true, would provide the best explanation of our evidence—is in fact true. Lewis's third rule is the Rule of Conservatism. We may, defeasibly, ignore possibilities that we know are commonly ignored by those around us.[30]

Although Lewis, in elaborating his definition, is sensitive to subtle details about the way the predicate "knows that" is commonly used, he does not explicitly say whether he intends his definition to be an analysis of what is actually meant by

---

[29] The qualification is important because *in conjunction with his final three rules*, it enables Lewis to rebut objections raised by such writers as Jonathan Vogel. Vogel (1999) described several troublesome possibilities that saliently resemble actuality so far as the subject's evidence is concerned but that cannot, as he saw it, be eliminated Lewis's rules. Vogel did not, however, at least in my opinion, take adequate account of Lewis's rules of Reliability, Method, and Conservatism, which can show that these possibilities deserve to be ignored in any normal context. (In fairness to Vogel I should add that in his appendix to the paper he expressed a cautious attitude toward his criticism, saying that it "should be taken as exploratory rather than final" (p. 172).

[30] These last three rules are described on pp. 242f of Lewis (1999).

the predicate, at least in some favored dialect, or a reconstruction of what is thus meant. I think it is obvious that his definition does not capture what is actually meant by most careful speakers of English. According to his definition, people know many things they have no conception of, for their evidence on numerous occasions fails to match the relevant alternatives to an actual possibility that they cannot actually comprehend. As an illustration of this, consider little Patty who is standing before a kangaroo in a zoo. I am not sure what the relevant alternative possibilities to an instance of seeing an actual kangaroo may be, but little Patty's entire perceptual experience and memory on this occasion may fail to match all of them. (The experience of perceiving an adult kangaroo in good light is not realistically similar to perceiving anything else that I can think of.) Yet little Patty may have no idea what a kangaroo is, even though perceiving such a thing does fit her perceptual experience. In this sort of case her evidence does eliminate the relative alternatives to seeing a kangaroo, but most discerning speakers would not agree that little Patty therefore knows she is perceiving a kangaroo. There are, of course, many things little Patty does know in this situation: she knows she is seeing a large furry animal with a large funny tail, for instance. But there is nothing in Lewis's definition, which is focused on possibilities matching perceptual experience and memory, that requires a person to understand (or comprehend) the possibility that his or her evidence *fails to eliminate*.

Lewis, true to my conception, at least, of routine cases of knowledge-that, insists that knowing does not require belief or even a justification the subject can give, but he overlooks a requirement that I would emphasize—namely, that a knower possess appropriate information. In speaking of information here I have in mind something propositional and true that one can mentally possess as the result of learning and then retain without necessarily believing that one possesses it or thinking that it is true. An example of information so understood is what I learned when I was taught, or discovered,[31] the expansion of π to five decimal places. One might think of this as knowledge, but I am thinking of it, perhaps idiosyncratically, as something more elemental, something one could possess unknowingly and without supporting evidence.[32] If I have learned that the decimal expansion of π to five decimal places is 3.14159, I may come to believe that I have forgotten it, have no belief about what it is, and no longer recall how or from whom I learned it. If I am urged to identify the sequence of integers defining it, I may nevertheless succeed in producing it and be surprised by my accomplishment, deciding that I did not forget it after all. If, generally speaking, I actually retain certain information, I must be able to produce it if I am called upon to do so or stimulated by some reminder. There is no doubt a fine line between being reminded and being taught in the way the slave boy in the *Meno* was taught a geometrical theorem by Socrates' questions, but it is clear that a person can genuinely possess information without realizing it and without being able to bring it to mind except by means of some information-eliciting reminder.

As I see it, then, we are prepared to say that someone knows that P only when we are convinced that he or she possesses the information that P[33] and that this information is adequately supported, though not necessarily insured, by appro-

[31] I am not sure how I gained this information. I probably obtained it from a teacher in middle school, but I have no doubt that I subsequently verified it by computations I carried out myself. I mention this because it is a simple example of an important phenomenon: We are generally uncertain about the source of much of our information, or the evidence we have for many of the things we say we know.

[32] My conception of information here is significantly different from the quantitative concept belonging to information theory that was introduced into epistemology by Dretske (1982).

[33] This conviction is also expressed by Bach (2005); see p. 63.

priate evidence possessed by that person.[34] I myself, like Lewis, would not require a knower to have some specific belief, but unlike Lewis I would require a knower to possess corresponding information in the sense I have described.[35] Such information is closely related to true belief, but it is not the same thing. Belief normally accompanies it, but not always.

Another shortcoming in Lewis's definition, if it is understood as an analysis of existing discourse, can be traced to his Rule of Attention. According to this rule, a possibility that is not ignored is not properly ignored. But real speakers, if they are self-confident, would insist on ignoring the possibility that the thing a child takes to be her kitty is really a robot, indistinguishable to sight and feeling from a real kitten, that aliens have perversely introduced. Instead of conceding that the child really doesn't know she has a kitty, they would normally dismiss this possibility as too far-fetched to be taken seriously. Norman Malcolm would have said that the child fails to know *for certain* that she has a kitten, but he would have no doubt that she knows it in a weak sense of "knows."[36] The mention of possibilities normally considered remote or far-fetched may make knowledge elusive if the hearer is someone with the sensibility of David Lewis, but not every sophisticated speaker of English would respond in the same way.[37]

The basic idea on which Lewis's definition rests--that a definite class of relevant possibilities is always pertinent to ascriptions of knowledge--is also very doubtful so far as existing usage is concerned. If we say that little Patty knows her name--knows that it is "Patty"--we need not have any alternatives in mind, other than the possibility that she does not know this;[38] and if I say that I know my neighbor's dog is an Airedale, my claim is unlikely to be assessed by ruling out some set of alternative possibilities, such as that it is a Scotty, a Welsh terrier, or a large mongrel with kinky hair. (No one familiar with Airedales would confuse one with any other animal.) We sometimes do have a number of alternative possibilities in mind when we assess a knowledge claim, but we do not always have a group of them in mind, and the ones that we might consider are not plausibly a function of our epistemic situation or that of someone ascribing knowledge to us. Fred Dretske, who originally introduced the alternative-possibilities idea into discussions of knowledge,[39] did not make a compelling case for the epistemic necessity of such alternatives, for he in effect used the idea of relevant alternatives to disambiguate a claim to knowledge. But if what someone might know is stated in unambiguous terms, no relative alternatives need to be mentioned. Thus, instead of clarifying the ambiguous "Lefty killed Otto" by saying "It was Lefty rather than George or Mike who killed Otto," one could say "Lefty was the person who killed Otto." And instead of saying, "Lefty killed Otto rather than injuring or threatening him," one could say, "What Lefty did to Otto was to kill him."[40]

---

[34] I am thinking of evidence here the way Lewis does: "the subject's entire perceptual experience and memory." Evidence so understood is a subject's total evidence at a time. See footnote 46 below.

[35] I therefore reject the widely accepted idea, defended by Plantinga, that knowing that P is having a true belief that P and a proper warrant for it. See Plantinga (1993).

[36] See above, p. 5.

[37] I discuss this further below. But see also Bach (2005), p. 86.

[38] This is my basic objection to the contrastivist position defended by Jonathan Schaffer (2005). Perhaps a more telling example against him is this. If I say "Tommy certainly knows what a sexual virgin is," I imply that Tommy knows that a sexual virgin is a person who has not engaged in sexual intercourse, but I do not imply that Tommy knows this rather than some other thing that I am or anyone else is apt to mention.

[39] See Fred Dretske (1970).

[40] Schaffer is well aware of this fact, but he holds a disjunctive view nevertheless. See Schaffer (2005), 251f.

If Lewis's definition is understood not as an analysis of an existing meaning but an "explication" or "rational reconstruction" of it, the objections I have been making do not apply--or may not apply, depending on how close to the vernacular such a reconstruction is expected to be.[41] I believe that Lewis' definition is not sufficiently close to be an acceptable reconstruction, at least as I understand the verb "knows," but his definition can be modified to bring it closer. An obvious way of doing this is to add an appropriate conjunct to his definition--specifically, that the subject understands and indeed has the information that P. The idea of having information in my sense is of course vague, but Lewis's definition, in spite of its surface crispness, is considerably vague already. This is evident in his formulation of the rules for identifying possibilities that may or may not be properly ignored in determining whether a subject knows that P in a particular context.

Although Lewis's reconstruction (for that is really what it is) can be brought closer to existing usage, it will remain unacceptably distant, at least for me, because of the alternative-possibility idea on which it is based. There are two difficulties I have in mind here. The first is that ascriptions of knowledge that would normally be regarded as inconsistent may turn out to be compatible on Lewis's reconstruction. If I say that Harry knows that his dog is safe at home and Mary denies this a little later, saying that for all Harry actually knows his dog may have got out of the house and taken to the streets, my claim and Mary's counter-claim would normally be regarded as inconsistent, one of us being right and the other wrong. But if our claims are, in effect, relativized to different sets of alternatives that Harry's evidence is supposed to rule out, they appear to be compatible, since "Harry's evidence rules the alternatives in set A" would seem to be consistent with "Harry's evidence rules out the alternatives in set B" if A is distinct from or not included in B. The second difficulty concerns the task of ruling out *all* the possibilities that are supposed to be relevant in a particular case. How can we be sure that we have done this? Are our everyday knowledge ascriptions really so precise in their implications that the relative alternatives form a determinate set? I should say no. Yet even if a definite set is assumed to exist in a particular case, Lewis's Rule of Attention may nevertheless frustrate our efforts to identify a single survivor, for it may generate further possibilities that become relevant merely because, wondering if they are relevant, we bring them to mind. Thus the question whether a subject's evidence is adequate for knowledge even in some humdrum case may prove to be persistently elusive for a conscientious and imaginative investigator.

A reconstructed concept is actually a new concept, not a clarified version of an old one; and if someone wants to adopt Lewis's concept, he or she may certainly do so in spite of the problems I think it involves. For my part, there is no need to do struggle with such problems; a dual conception of knowing that is more in line with existing usage can actually be provided. The conception I shall develop is not a form of contextualism or epistemic relativism, but theories of those kinds other than Lewis's have thus far failed to show how the stronger or weaker standards of assessment that they speak of are to be identified in a given context and how a given standard can be shown to be correct. Until this is done, a contextualist theory not featuring the relevant-alternatives idea that I find objectionable is not preferable to the theory that Lewis developed. To my mind, his is the best of its kind.[42]

---

[41] The notion of an "explication" or "rational reconstruction" was central to the analytical work of Rudolf Carnap; I will discuss it further in Ch. 3. See Carnap (1956), pp. 7f.

[42] Feldman (1999), p. 109, complains that Lewis is committed to the view that a person's evidence never eliminates a possibility on theoretical grounds having to do with theory choice and the like. But Lewis's Rules of Method may justify us in ignoring possibilities on such grounds, at least if we don't explicitly consider those possibilities. As my criticism of Lewis above makes clear, I would agree with Feldman that Lewis's Rule of Attention is excessively demanding and should be revised.

A Dual Analysis of Knowledge

As I see it, two different senses of "knows that" are compatible with existing philosophical usage. One is that of knowing for certain, a sense possibly introduced into English by philosophers in the tradition of Plato and Descartes. One who has certain knowledge of some fact must have conclusive evidence that the corresponding proposition is true. In the strictest sense this proposition is, for the knower,[43] either immediately certain (a perceptual certainty or a trivial analytic truth, as empiricists might say) or provable by a sequence of elementary valid inferences proceeding, ultimately, from immediately certain premises.[44] In a weaker sense suggested by Hume, the proposition may be roughly described as provable, for the knower, by "such arguments from experience as leave no room for doubt or opposition."[45] As I noted, the scope of certain knowledge is significantly limited, but it does not raise the kind of problems featured in the literature consequent to Gettier's critical paper, and a means of avoiding those problems need not be incorporated into a satisfactory account of such knowledge.

The other sense of "knows that" compatible with current philosophical usage does not require rational certainty in either of the forms I have just described. This sense purports to be non-technical, and it is less determinate than a philosopher might wish. Those who worry about the analysis of knowledge generally agree that if S knows that P, it is true that P, and, like Gettier, they generally suppose that S may know that P even in cases where S lacks evidence sufficient to insure that P.[46] In other respects they often disagree about the implications of this sense of the expression. Most still assume that if S knows that P, S believes that P and has some justification for this belief. As I noted, David Lewis rejects both implications; he thinks that reflection on actual cases makes it inadvisable to accept either one. If a large number of philosophically innocent speakers were asked whether Lewis were right in rejecting these implications, they would probably disagree among themselves, depending on who asks the question and how. The conclusion I draw from this likely disagreement is that common speech is not precise and ordinary speakers do not make the distinctions that philosophers, reflecting on that speech, are apt to draw. Thus, because speakers who know that P are commonly aware of what they think they know, it is usually true that S believes that P when S knows that P; as a result, the possibility of knowing that P and not believing it is not normally considered. If it is brought up, different speakers may respond differently and so become aware of differences in the meanings they attach to the predicate "knows."

Reflection on possible cases has induced me, as it did Lewis, to hold that knowing that P (in the weak sense) should not imply either believing that P or having a justification for so believing. As I conceive of such knowing, one who knows that P must have the information that P and must have, or at least have had,[47] good evidence that P,[48] but this evidence need not be irreducibly propositional: it may feature

---

[43] By saying "for the knower" here I mean that the knower is able to give the proof and is aware that he (or she) can do so.

[44] The immediate certainty of the proposition or the acceptability of the proof comprises the knower's *basis* for this knowledge. I discuss the notions of certainty and proof in later chapters. What I say here should be understood as provisional.

[45] Hume, *Inquiry Concerning Human Understanding*, VI, 46.

[46] As I noted, Fogelin (1994) and Williamson (2000) among others apparently disagree with this.

[47] See footnote 78 below.

[48] The subject's *total evidence* must adequately support the proposition that P; he or she must not have some negative evidence that effectively refutes or casts doubt on this proposition.

a condition or response that renders the proposition that P highly probable in the circumstances, worthy of acting on by the knower or someone apprised of that evidence. Thus, people familiar with North American lakes may know that a birdcall they hear is that of a loon, even though they cannot say how they know this. They simply know what a loon sounds like. Such people are similar to violinists who can tell and thus know when their instrument is in tune by hearing how it sounds when they play it.[49] People who know something on a similar basis—the basis of how a thing looks, sounds, feels, or tastes—may of course be able, sometimes, to provide a justification for what they know, saying "I know that P because I judge that P and my past behavior shows I have the capacity to make true judgments of this kind spontaneously." But knowers need not be skilled in providing justifications; if they did, most ordinary people would not be knowers even if they had a Socrates available to help them construct a justification.

Cases of knowledge not based on the experience of how something looks, sounds, smells, tastes, or feels may be based on propositional evidence—on all the presumptions and inferential methods Lewis mentions in addition to memory evidence: the testimony of reputable observers, the received opinion of scientific authorities, generalizations from test results, and possibly even inferences to the best explanation.[50] Here knowledge is typically based on further knowledge, although there is no general requirement to this effect. Some knowledge is simply more basic than other knowledge. As far as evidence *per se* is concerned, the most general claim one can reasonably make about its quality is that it must be "good" if it supports knowledge. A more informative way of putting this, perhaps, is to say that if S knows that P, the probability of P on the basis of S's total evidence must be, in the circumstances, high enough to convince an appropriately informed, clear-headed, and impartial observer that S has the relevant information.[51] Being "appropriately" informed involves being familiar with relevant standards of assessment and having the pertinent discriminative skills (think of a judge at a dog show); being clear-headed and impartial are qualities well enough understood not to require elaboration here. As Lewis in effect emphasized, ordinary knowledge assessments are based on a fund of common experience and inferential methods that are normally regarded as acceptable but are, as Lewis put it, "defeasible." These assessments are sufficient for common assessments of knowledge. In cases of rational certainty, the probability of P on the basis of the subject's evidence is 1 (or maximal); if P is immediately certain the probability of P is maximal on the basis of its intrinsic character. But more on this later.

When people ascribe knowledge to others in everyday life, they generally have no doubt that others have the appropriate information, but they are often very casual about the character (the strength or quality) of the evidence these people

---

[49] As I emphasize later in chapter five, people who know something on the basis of how a thing looks, sounds, feels, or the like must normally have further knowledge: a man who knows he is hearing a loon must know what a loon is, for instance. This further knowledge need not function as premise in a possible inference supporting the proposition known, however. It is required for having the kind of information needed for knowledge in this case.

[50] In chapter six I criticize this widely accepted form of a posteriori inference, but I defend another form of inference as a substitute for it. As far as evidence itself is concerned, note that one can be rationally confident that S has good evidence for P without knowing exactly what that evidence is. Think of the evidence a normal American adult has for the sort of animal a monkey is or of (J.L. Austin's example) "I was brought up in the fens" as an answer to the question, "How do you know that is a bittern?" On this last matter, see footnote 78.

[51] I discuss inductive (or evidential) probability in some detail in chapter six.

possess.[52] They suppose that a person's evidence is good in the circumstances; they simply have a fairly hazy idea of what that evidence is. I do not believe that acceptable standards for the quality of a subject's evidence vary in the way contextualists suppose, but I have no doubt that stronger evidence is normally demanded when the consequences of acting on erroneous input are uncertain, hazardous, or incompatible, in some significant way, with the agent's or the observer's purposes,[53] or when assumptions customarily used in assessing knowledge claims are put in question by some unusual fact or circumstance. In mentioning these instances I do not want to suggest that I am aware of some principle by which one can invariably identify conditions where strict standards are "appropriate." As far as I can tell, individual persons generally call the tune here. If they are convinced that a knowledge claim that would normally be accepted in the circumstances is doubtful in a way they consider significant, they will generally demand higher standards of evidence.[54] But significant doubt is often an arguable matter, not easily resolved. Some doubts may be considered neurotic; others may be considered idle or even "merely philosophical." In later chapters I shall discuss examples in which ordinary standards are questioned for plausible philosophical reasons: they are claimed to be insufficiently critical.

I have said that a subject who knows that P must possess the information that P.[55] A condition the information that P must meet if S knows that P is that S should possess it because of some kind of learning or evidence. This condition is emphasized in good English dictionaries. According to the new *Oxford American Desk Dictionary and Thesaurus,* "**Knowledge** applies to any body of fact gathered by study, observation, or experience, and to the ideas inferred from these facts…."[56] This assertion is not presented in the way a philosopher would put it, but the idea is clear enough: knowledge that P results from a reliable process of fact-gathering or fact-assimilation, one sufficient to render P sufficiently probable in the circumstances to satisfy the condition I mentioned earlier.[57] To have obtained a fact (something true) by such a process is to possess information. It is analytic of the idea of information, at least as I understand it, that it is ultimately created this way. It may, as I implied earlier, be retained when the supporting evidence is lost, and it may even be transmitted to another person on a responsible subject's authority. But if a thought or supposition that happens to be true simply comes to mind and has no evidential support that anyone is aware of, it is not properly information in the relevant sense.

---

[52] Kent Bach (2005) recently expressed this idea particularly well, saying "It is worth keeping in mind that most of the time, outside philosophy, when we consider whether someone knows something, we are mainly interested in whether the person has the information, not in whether the person's belief rises to the level of knowledge. Ordinarily we do not already assume that they have a true belief and just focus on whether their epistemic position suffices for knowledge. Similarly, when we say that someone does not know something, we typically mean that they don't have the information" (pp. 62f).

[53] Jason Stanley (2006) constructed an interest-relative invariantist account of knowledge assessment largely on the basis such considerations; for critical remarks on his view, which I consider compelling, see Schiffer, "Interest-Relative Invariantism," (forthcoming).

[54] There is an extensive literature concerned with the relation between knowing and what some might consider remote possibilities of error. This literature and the subject it concerns are admirably discussed in Hawthorne (2004).

[55] The condition on information that I mention here applies to knowing "for certain" as well as to knowing in the weaker sense. In what follows I shall use "knowing" or "knowledge" to refer to the weak sense of these terms; I shall add a qualifier when referring to the stronger sense.

[56] *Oxford American Desk Dictionary and Thesaurus,* p. 460.

[57] I discuss probability and rational acceptability in chapter six.

Problems for Two Senses of "Knows that P"

There is a lot more to say about the senses of "knows" I have described, but before proceeding with it I must consider an important objection to the idea that these senses are actually compatible with reflective speech about knowledge. The objection arises from a verbal phenomenon emphasized by John Macfarlane in a recent paper.[58] Although we do often require stronger or weaker evidence in accepting knowledge claims or ascriptions, if we first say we know something P on the basis of contextually acceptable evidence but then, in view of facts brought to our attention, deny that we know it because stronger evidence now seems required, we will not generally allow that we did, and do, know P in a weak sense but do not know it in a stronger one. We will deny that we ever knew P at all. Macfarlane illustrates this fact by the following example. Having left my car in my driveway when I go to work, I later claim that I know where my car is: it is in my driveway at home. Even though I may believe that my neighborhood is a safe one where thefts rarely occur, I may nevertheless concede, when the possibility of theft is strongly emphasized, that I do not really know what I said I knew. In saying this, I am now using a stronger epistemic standard than what I used when I originally made my claim, but I do not insist that I still know in a weak sense what I now deny on the basis of stronger evidence. If there really were stronger and weaker senses of 'know," the fact that I do not know something in a strong sense should not prompt me to deny that know it in a weaker sense. In the case presented I clearly change my mind, something that I would not be doing if I merely made claims using two different senses of the same word.

The example is impressive, but it does not actually undermine my two senses view. Here is why. In both senses of the verb, "S knows that P" implies that "P" is true. When I made the ascription on weaker evidence, I had no doubt that "P" was true; I was wholly confident that it was true, although my evidence fell far short of what is required for rational certainty. But the facts that prompted a stronger epistemic assessment of my knowledge ascription rendered the truth of "P" doubtful; I not only denied knowing "P" on a stronger evidence base, but I came to doubt that "P" is actually true. This doubt undermined my initial claim as well as any temptation I may have felt to claim knowledge of "P" in a weaker sense.

Another reason why the example fails to undermine the two-sense view is that the two cases presented are both involve a weak sense of "knows." Weightier evidence is brought to bear on the second case, but different cases of imperfect knowledge may well involve evidence of different weights. When I take more evidence into account, I change my mind about what I know; I do not suppose I am moving to a different sense of "knows." The case Macfarlane describes is not applicable, therefore, two senses of "knows" I have described. If I were to say that I know my car is safer on stronger evidence, I could still deny that I knew it for certain—that I could actually prove that it is true. Knowing matters of fact in this way is something that may elude our powers. I shall return to this matter in the last part of my last chapter.

Although the senses I have distinguished are not undermined by the phenomenon Macfarlane identified, one or the other might seem threatened by a problem routinely posed in discussions of skepticism. The problem is sometimes called the closure paradox,[59] and it may be stated means of the following three assertions:

1. Moore knows that he has hands.

---

[58] See Macfarlane (2005). The phenomenon also provides an objection to contextualism that is closely related to the first one I give on p. 9.

[59] See DeRose (1995) and e.g. Schaffer (2005), pp. 259f.

2. Moore doesn't know he is not a brain in a vat.
3. If Moore doesn't know he is not a brain in a vat, then he doesn't know that he has hands.

These assertions are thought to be individually plausible but jointly contradictory. How is the paradox to be resolved?

Dogmatists typically accept (1) and (3) but deny (2); skeptics accept (2) and (3) but deny (1); and contextualists accept (1) and (2) but deny (3). Given my two senses of "knows," I can consistently accept both (1) and (2), saying that Moore knows he has hands in the weak sense, the sense presupposing conventional assumptions about perception, but fails to know *for certain* that he is not a brain in a vat. But if I accept both (1) and (2), I must deny (3), a move that appears to deny "closure" for known logical implication.[60] Jonathan Schaffer contends that this is a serious blunder on two counts: it "seems absurd" on the face of it and it "collapses inferences." Deduction surely transmits knowledge, he says, for mathematical proof is based on deductive inference and it unquestionably yields knowledge.[61] Am I really making a blunder if I deny (3)?

The answer is no. Closure holds only for logical or mathematical operations denoted by univocal predicates. If I know that P and also know (in the same sense of the word) that P entails Q, I certainly know that Q (in the same sense of "know").[62] But if I know that P in a weak sense and know that P entails Q in the same or even a stronger sense, closure does not require that I know that Q in the strong sense assumed by premise (2). Anyone who accepts a weak sense of "know" should allow that Moore knows that he is not a disembodied brain in a vat in the weak sense in which he knows that he has hands. How could he be disembodied if he has hands? And how could he know that he is disembodied without knowing this? Having this knowledge is not knowing for certain, of course; it is not knowing in a way that allows no possibility of error. Moore, a particularly self-confident person, might conceivably have believed that he had certain knowledge for the proposition that he has hands, but if he knew this for certain, he would equally know for certain that he is not disembodied, not something that doesn't have any physical part or appendage. The so-called paradox of closure therefore falls apart on examination, in my opinion. Statements (1) and (2) are jointly acceptable only when they include different senses of "knows." But then, when they contain these different senses, statement (3) does not involve a closure principle. For a univocal sense of "knows," a closure principle (suitably qualified)[63] is, I believe, patently acceptable.

Avoiding Gettier Counterexamples

Although Gettier counter-examples are not, as I have argued, pertinent to cases of certain knowledge, they do apply to knowledge in a weak sense, a subject's evidence for which is logically compatible with the falsity of what he or she is said to know. Any evidence that is sufficient for knowledge in this sense must satisfy a condition

---

[60] Closure in this case is the principle commonly formulated by saying that if S knows that P and also knows that P entails Q, S knows that Q. To avoid certain problems not pertinent to the issue here, Hawthorne (2005) formulates the principle more cautiously, saying that if S knows that P and competently deduces Q from P, thereby coming to believe Q while retaining his knowledge that P, S comes to know that Q.

[61] See Schaffer (2005), p. 261.

[62] This is denied by Dretske (2005) mainly on the ground that a subject's means of knowing that P is not inevitably "transmitted" to Q even when he or she knows that P entails Q. But subjects having the latter knowledge have an extra premise to use in inferring Q, one that makes explicit something involved in knowing that P. Their basis for accepting Q is therefore possibly different from their means of knowing that P.

[63] See footnote 58.

that rules out Gettier's counter-examples and others relevantly like them. The condition is needed because thoughtful speakers of English who do not require conclusive evidence for knowledge uniformly deny that Gettier examples are genuine cases of knowing, and their conception of knowing must include some basis for this negative attitude. The relevant condition has proved notoriously difficult to pin down in a satisfactory way, however. It is expressed vaguely by saying that the truth of the proposition embodying the information must not be "accidental" so far as that evidence is concerned.[64] The fact that P is strongly supported by evidence E does not insure that P is not accidental in the relevant sense. In Gettier's first example, Smith has good evidence for the disjunction "Jones owns a Ford or Brown is in Barcelona" because he has good evidence for the first disjunct, "Jones owns a Ford," and knows that the disjunction is a logical consequence of it. But the truth of the disjunction depends on the truth of the second disjunct, "Brown is in Barcelona," to which Smith's evidence is not pertinent at all. Gettier's second example is similar. Although Smith has good evidence for the conclusion that the man who will get the job has ten coins in his pocket, his evidence pertains entirely to Jones and has nothing specifically to do with the person who will actually get the job and who actually has the coins in his pocket. He has good reason to believe that Jones will get the job and that Jones has ten coins in his pocket, but the truth-maker for the conclusion is a compound fact about Smith himself, something that is purely accidental so far Smith's evidence is concerned.

The vagueness of referring to a truth that is "accidental" in relation to given evidence is not entirely damning for an explanation of what is meant by a vernacular expression, which can be expected to share that vagueness, but I think it is possible to specify the relevant condition in philosophically more congenial terms. The basic idea, which leads ordinary speakers of English to reject Gettier cases as genuine cases of knowledge, is that the conclusions formed in those cases do not qualify as expressions of knowledge because they are not *made true* by facts for which the subject possesses evidence. (If S possesses evidence for a particular fact, I shall say that the fact is evidentially accessible to S.)

The notion of making true that I am employing here is frequently used in truth-conditional semantics.[65] Elementary statements are made true by pertinent facts about the reality they concern; formulas of the form $\sim\Phi$ are made true by the fact that the inner formula, Φ, is not true; conjunctions are made true by the fact that both conjuncts are true; disjunctions are made true by the fact that one or the other of their disjuncts is true; material conditionals are made true by the fact that either their antecedents are false or their consequents are true; quantified formulas are made true, ultimately, by pertinent facts about individual entities, either all or some, in the appropriate domain of quantification; and modal facts are made true by facts about possible worlds or the contents of pertinent concepts. This idea clearly applies to the two Gettier examples I mentioned. The statement that either Jones owns a Ford or Brown is in Barcelona is made true, ultimately, by the fact that Brown is in Barcelona; and the statement that the man who will get the relevant job has ten coins in his pocket is made true by the facts that Smith is the one who will get that job and that Smith has ten coins in his pocket.

---

[64] The last part of this condition is widely recognized; for example, Steup (1996) expresses something like it by means of the expressions "lucky truth" and "lucky guess" (see, p. 9), and Heller (1999) describes his preferred theory of knowledge as an "anti-luck" theory.

[65] I am not employing the notion of a truth-maker recently worked out by Armstrong (2004); I do not believe that such an elaborate notion is needed for my purposes here.

The idea of a truth-maker, as I am using it, is reasonably unproblematic,[66] but my notion of evidential access requires more discussion. Consider the following example. Living in a rural area, I discover some curious scratches on the door of a garden shed, and I ask my neighbor about them. They look as if they were made by the teeth of some animal, and my neighbor is far more knowledgeable about the local wildlife than I am. He says that a fox or a raccoon made the scratches. Suppose a fox actually made them and that my neighbor's experience provides good evidence that they were made by a fox or a raccoon. Off-hand, it would appear that my neighbor knows, in the weak sense, that one or the other of these animals did cause the scratches, but does he have evidential access to the truth-maker for the disjunction involved here? He certainly cannot say what that truth-maker is.

According to my stipulation regarding evidential access, a person who has evidential access to a truth-maker need not know what that truth-maker is; it is sufficient that he have evidence for it. In the case in question, the neighbor has such evidence, but he also has evidence supporting another hypothesis. Primarily, his confidence is attached to a disjunction: his evidence concerns both disjuncts equally well. No one would think of this case as a Gettier example, so there is no perception (on the assumption that the man's judgment is good and his experience is extensive) that he is right because of a lucky guess. He judges that a disjunction is true, and his evidence supports both disjuncts sufficiently well to make him unwilling to detach one in favor of the other. If his evidence favored just one of the disjuncts, he would have knowledge of the disjunction only if his evidence favored the truth-maker in a sufficiently strong way. But this is not the way the case is specified.

In view of objections I have heard, I must emphasize that the evidence favoring a truth-maker for a compound proposition need not be sufficient to justify belief in that truth-maker. This is evident from the case of the scratches on the door. The neighbor has good evidence that a fox or a raccoon made the scratches, but his evidence does not warrant his believing that a fox made them or his believing that a raccoon made them. Since these simple hypotheses are incompatible, the probability of their disjunction on the basis of his evidence—that is, Prob(F or R, on E)—is equal to the probability of the fox hypothesis on E plus the probability of the raccoon hypothesis on E—that is, to Prob(F on E) + Prob(R on E). But although the fact that the conditional probability of the disjunction is high enough to support the conclusion that the neighbor knows that disjunction is true, it does not follow that the conditional probability of either simple hypothesis high enough to support the conclusion that it is true. In the case I have described these latter hypotheses are too weakly supported for such a conclusion. On the other hand, the support that the disjunction receives from the evidence is owing to the non-negligible conditional probabilities of these simple hypotheses. If the sum of these simple probabilities were sufficiently low, the conditional probability of the disjunction would be too low for knowledge.

What about logical truths? If a person knows that it will rain tomorrow or that it will not, need he have evidence for the contingency that happens to make this disjunction true? The answer is no. My truth-maker requirement holds only for contingent knowledge. A tautology is true no matter what the contingent facts may be, so knowledge of its truth does not require evidence for such facts. The aim of the requirement is to rule out Gettier cases, and these cases do not concern tautologies or analytic truths.

---

[66] The reader should realize that the clauses I give in the last paragraph amount to a recursive characterization of a truth-maker. A disjunctive statement is true if at least one disjunct is true, but to ascertain the truth-maker for one of the disjuncts, we must attend to the logical structure of that disjunct. If, for example, it is an existentially quantified formula, its truth-maker must be sought by means of another clause, the one governing quantified formulas.

It is obvious that if we can know (in the weak sense) on the basis of memory, testimony, or experimental inference, our evidential access to a truth-maker can sometimes be considerably indirect. It must, however, exist if we have genuine knowledge about a contingent matter. My conception of this access is far from precise, but most everyday concepts share this vagueness. We can recognize clear cases where this access is lacking: if we could not do so, we could not recognize a Gettier example when we are presented with one. We can also recognize clear cases where this kind of access is present; I believe I have cited one already. There are bound to be borderline cases where we cannot say whether it is present or not, but this is characteristic of everyday concepts. With them, vagueness is the rule rather than the exception.

I should emphasize at this point that a subject's evidential access to a truth-maker does not generally provide all the information that a knower needs. This is evident from what is obviously involved in knowing something disjunctive. If I know that P or Q, then the truth-maker for my belief, if the belief is contingent and "P" and "Q" are not themselves compound or quantified propositions, is either the fact that P or the fact that Q.[67] But if my belief is made true by the fact that P, I will not have the information that P or Q unless I realize that (P or Q) follows from P. As a general matter, evidence for the truth-maker of Φ will provide adequate evidence for Φ, but further evidence may be required for the information condition.

Another, more interesting illustration is the phony barn case that is well known in the literature on knowing.[68] This case is often described as a Gettier example, but it is not sufficiently similar to the examples Gettier actually gave to merit that description. It can be set down as follows. Riding on a train through a section of countryside that contains, unknown to me, numerous facsimiles of barn facades, I have a clear perception of a real barn and thus have evidential access, by means of perception, to the truth-maker for my belief. But I do not know that I am seeing a barn: my evidence is not, in the circumstances, good enough. It is the sort of perceptual evidence that would ordinarily be sufficient for knowing in a sense not requiring rational certainty, but it is not sufficient in the specified case because of the presence of the barn-facade facsimiles, which I could not distinguish from the real thing at a distance and through the window of a train. This case does not require me to adopt a strategy of ruling out relevant alternatives, although it would be reasonable to do so. What I need is better access to the object I perceive. The normal, defeasible presumption that what I seem to be seeing in the light of day is what is actually there is undermined by the presence of the facsimiles, and I therefore need more evidence than what is normally sufficient for knowing what I am seeing. What I positively need is to make a closer inspection of the object; I must ascertain what actually possesses the façade that I discern through the window of the train.

There are other examples, originally offered as ostensible counter-instances to a traditional conception of knowing and sometimes considered Gettier cases, that are similar to this last one. A lecturer may look at a clock that is normally reliable and form the belief that the time is 4:15 p.m.; the clock is out of order, but its hands

---

[67] If both P and Q are true, both P and Q make the disjunction true. Should someone who knows that the disjunction is true possess evidence for both disjuncts in this case? The answer is no. We have a redundant truth-maker here. Since the truth of either disjunct insures the truth of the disjunction, evidence for either is sufficient for knowing. An analogous point holds for the knowledge of existentially quantified truths. If four different books are on a table, evidence that any one of the books is there is enough to satisfy the truth-maker requirement for "At least one book is on the table." When redundant truth-makers are available, evidence for a *sufficient* truth-maker satisfies the truth-maker condition.

[68] Lewis discusses this case, originally described by Carl Ginet according to Goldman (1976), in connection with Gettier examples. I myself do not consider the case a genuine Gettier example because it does not conform to the pattern of the examples Gettier actually gave. His examples were distinguished by the lack of evidential access to the pertinent truth-maker, which is not present in the phony barn case.

happen to indicate that it is 4:15 p.m., and this happens to be the right time. Or again a man wishing to deceive a certain person visiting another country prepares a fake front page of a reliable newspaper bearing the headline, "REVOLUTION IN BRAZIL!" On reading the contrived news story, the victim forms the belief that such a revolution has occurred, and by an extraordinary coincidence his belief turns out to be true—to the utter surprise of the deceiver. Do these cases satisfy the definition of knowing, or not? They do not. They do not satisfy my definition for the same reason that the phony barn case does not satisfy it: the person's evidence is not adequate in the circumstances. The clock and a newspaper whose front page has the appearance of the fake one are both generally reliable: what they indicate is normally true and normally an adequate basis for knowing in a sense not requiring rational certainty. But these are not normal circumstances and in these circumstances the defeasible[69] evidence that the clock and the paper provide is not good enough for knowing. It is defeated by the abnormality of the circumstances, and it does not provide genuine information about the correct time and the occurrence of the revolution.

Concluding Remarks

I have distinguished two senses of "knows that" in this chapter, one involving rational certainty and the other not. The basis for my distinction was the common practice of requiring rational certainty in some cases but not requiring it in others. Sometimes people are asked "Do you know that for certain?" and sometimes this question never arises.

A skeptic might insist that there is just one sense of "knows," a strong one, which is sometimes loosely applied to cases that do not really deserve to be called knowledge at all. But this opinion is clearly at odds with actual usage. As I noted, philosophers have for centuries associated knowing with being rationally certain about something, but ordinary people employ a much looser and less strict idea. If you tell tough-minded farmers or gardeners that they do not really know they are holding a shovel or pitchfork because they *might* actually be dreaming or hallucinating, they will simply laugh or roll their eyes, and walk away. They have no doubt that these possibilities are fanciful. Their opinion on this matter is not apt to be challenged by someone with scientific interests. There is surely some scientific knowledge, but no one seriously concerned with empirical matters would want to contend that matters of fact are self-evident or capable of proof.

Since it is philosophers rather than farmers, bakers, or laboratory workers who typically insist on a rational certainty requirement, we might adopt Bishop Butler's language and call one sense the "strict and philosophical" one and the other the "loose and popular."[70] I have acknowledged that in real life the quality of evidence needed for agreement on knowledge ascriptions tends to be higher in contexts where the consequences of acting on erroneous input are considered uncertain, hazardous, or incompatible, in some significant way, with the agent's or the observer's purposes, or where assumptions customarily used in assessing knowledge claims are put in question by some unusual fact or circumstance. But this kind of variation, which is much less extensive than Lewis's account implies, is not incompatible with just two senses of "knows." In most of the contexts where stricter standards are insisted upon, rational certainty as a philosopher would understand it is not actually

---

[69] See p. 12 above.

[70] Chisholm (1976), p. 92, introduced this language into recent philosophy. His source was Butler (1839), pp. 263-70.

called for. The standards are strict, but they do not require logically tight demonstrations or perceptions of certain truth. They usually fall far short of what is demanded by a philosopher's "strict and philosophical" sense.

As a way of concluding this chapter, I want to bring together my analytical remarks about knowing by providing two pertinent definitions. In discussing what we might call the loose and popular sense of knowing that P, I said that a person having such knowledge must have the information that P but need not actually believe that P. I doubt that this qualification is advisable for the strict and philosophical sense. If you are rationally certain that P, it is hard to see how you might be unaware of what you know; in fact, you should be convinced that P is true.[71] In view of this the following definition, which recalls the familiar justified-true-belief definition, seems satisfactory for the stricter, philosophical sense:

> S knows that P for certain just when (i) it is true that P, (ii) S is rationally convinced that P, and (iii) S's rational conviction (which involves a strong belief) that P is owing to S's awareness of evidence E that is conclusive for P: the probability of P on E is maximal, or 1.[72]

The nature of being rationally convinced by conclusive evidence is best left open here, because it is a serious source of disagreement. Both rationalists and empiricists will agree that being convinced of P this way amounts to having a proof for P or having some direct assurance that P is true, but they will disagree about what direct assurance might be. I will investigate this matter thoroughly in chapters two and three.

A definition for the kind of knowing that does not require rational certainty will be looser than the last one. I said earlier that the quality of the evidence required for non-certain (or, as I shall also call it, "imperfect") knowing must be good enough to convince an appropriately informed, clear-headed, and impartial referee that the subject has the relevant information. It does not have to rule out alternative possibilities in the way Lewis thought, because a specific set of such alternatives is not always available: what is said is generally not that determinate. So if, instead of trying to spell out the common presumptions and inferential methods that, in conjunction with the experiences and memories available to the subject, render the probability of the relevant proposition sufficiently high to be convincing in the way I said, we simply speak of the subject's evidence as being good in relation to that proposition and what is acceptably presumed about it, the subject, and the pertinent circumstances, we can say the following:

> S has "imperfect" knowledge that P just when (i) it is true that P; (ii) S has the information that P; (iii) S's evidence for P[73] is very good in relation to P and what is rationally presumed about it, the subject, and the pertinent circumstances; and (iv) S has evidential access to a sufficient truth-maker for P.[74]

This definition, which, like the former definition, omits a needed variable for the time at which the subject knows that P and abstracts from issues related to the context in

---

[71] If you have a true belief based on good evidence, you will have the information that P in the sense I have explained.

[72] In chapter six I discuss probability values of this kind in some detail. In this and the following definition "the proposition that" should be understood as preceding the last two occurrences of "P." I suppressed these words to simplify the formulas.

[73] If we allow that a subject may still have an item of information when he has forgotten its evidential source, the subject's evidence for P need not be possessed whenever he knows P. The same holds for his evidential access to a sufficient truth-maker.

[74] If the truth of P is a logical consequence of a more elementary truth-maker, any evidence satisfying clause (iv) will satisfy clause (iii).

which the knowledge that P is ascribed to the subject,[75] is still not as succinct as it could be, since it contains an obvious redundancy: clause (i) follows from clause (ii). But the provisions of a succinct definition are less easy to grasp than those of a more verbose one, and I am interested in being as helpful as I can.[76] Some of the key words in the definition have the meanings I specified in my earlier discussion. The words "the information that P" apply, I said, to something propositional and true that one can mentally possess as the result of learning and then retain without necessarily believing that one possesses it or thinking that it is true. And "evidential access to a truth-maker" applies to good evidence (not necessarily conclusive) for a fact that makes a certain proposition true.

As my discussion of Lewis's views on knowing made evident, routine ascriptions of knowledge are made on the presumption that the logically inconclusive evidence available to the subject is, in the circumstances, a sufficient indication of what the subject is being said to know. This presumption and others that I have mentioned are commonly made and, in practice, rarely questioned. When they are questioned, philosophical problems sometimes arise because it is difficult to show that the presumptions are in fact true and deserving of the confidence that is commonly placed upon them. If these problems cannot be solved, a skeptical reassessment of the supposed knowledge is apt to be made. The reassessment is important, whether we are happy with the concept of loose and popular knowledge or not. I shall discuss salient examples of these problems in chapters five and six.

[75] See the "general matter" I mention on p. 4 above.

[76] The reader might observe that this definition comes fairly close to the traditional one that Gettier criticized. In place of a belief requirement, it has an information requirement that attributes "something propositional and true" to a knower. This propositional attitude is kind of an attenuation of a belief requirement that also satisfies a truth condition. The evidence condition corresponds to the traditional justification condition, although it is logically weaker. What is entirely new is the access-to-the-truth-maker condition, which is needed to rule out Gettier cases.

Chapter 2
THE CLAIMS OF RATIONALISM

According to tradition, a fundamentally important kind of knowledge can be attained a priori, that is, independently of sensory experience. Knowledge that is independent of sensory experience in the required way may possibly require some sensory experience to obtain the ideas (or concepts) needed to comprehend the relevant proposition, but this experience would not be sufficient to ascertain that proposition's truth. Epistemological rationalists[1] give one characteristic account of how a priori knowledge is obtained and how it can apply to domains to which we lack experiential access. I shall discuss this account in the present chapter. Another account, the one I am committed on the whole to defending,[2] is given by logical empiricists, philosophers holding a twentieth-century version of the epistemological doctrine originally espoused by David Hume. I shall discuss this kind of empiricism in the chapter to follow.

The A Priori, Universality, and Necessity

If knowledge does not require rational certainty, anything that we can know a priori can also be known a posteriori: it can be obtained from experience.[3] Logical and mathematical knowledge is almost universally regarded as attainable a priori, by the use of reason itself, but it can also be learned from others, from parents or teachers, and accepted as knowledge on their authority. Most elementary logic and mathematics is in fact learned this way. But according to tradition, anything that can be known a priori must ultimately, if it is known at all, be known a priori—by someone, somewhere. The reason given for this is that a priori knowledge is universal and necessary, and nothing universal and necessary can be known first-hand on the basis of sense-experience.[4]

Some standard examples will illustrate why a priori knowledge is plausibly regarded as universal and necessary. Consider "All bodies are extended, or spread out in space." This judgment identifies a defining characteristic of a physical body; it specifies one feature that something must have if it is to count as such a body. A judgment of this kind is clearly universal in scope, since it holds for all physical bodies, wherever they may be and whenever they may exist. It is also necessary, since a thing cannot fail to have its defining characteristics. Or consider "2 + 3 = 5." This statement is universal in scope because it holds for all couples and all triples: any couple and any triple sums to a group of five. It also asserts something necessary, since no couple and no triple could fail to sum to a group of five.

A proposition that is universal and necessary could not be known by experience, Kant thought, because experience teaches us only that a thing is so and so,

---

[1] The classification "epistemological rationalist" is fairly loose. The term is commonly applied to philosophers holding the views I attribute to rationalists in this chapter, but those philosophers do not agree on all epistemological issues. A precise classification is not worth attempting, in my opinion.

[2] As I shall explain later in this chapter, I do not insist that all a priori knowledge *must* be analytic. I allow that some mathematical truths may not have this status, but if they do ultimately lack it, I have no idea what their ultimate justification is. Rationalist accounts of their truth, at least the kinds known to me, are unpersuasive—as I argue here. I consider myself a *moderate* rather than an extreme or *doctrinaire* logical empiricist.

[3] Perhaps a distinction is required here. If I know that P on a teacher's authority, I know it on the teacher's say-so and thus know it a posteriori. But what I thus come to know is nevertheless an a priori truth, something knowable a priori, and I may know, a posteriori, that it has this status.

[4] Kant gives this reason in his Introduction to the *Critique of Pure Reason*; see B3-B4.

not that it cannot be otherwise. Experience does justify us in making general statements such as "All bodies are heavy," but these statements are not "true and strict," Kant said, because their support is merely inductive: "We can properly only say that…so far as we have observed, there is no exception to this or that rule."[5] A teacher might convince us that some mathematical theorem is true and we might justifiably accept it on that teacher's authority, but we could not claim to know with certainty that it is true. To have that kind of knowledge, Kant thought, we would have to have first-hand knowledge of the relevant mathematical proof. Only a proof of this kind could assure us that the truth in question holds both universally and necessarily.

The view of a priori knowledge that I have been describing, which can be called the traditional view, is controversial today. W. V. O. Quine expressed the most general doubt about it in his famous paper, "Two Dogmas of Empiricism."[6] Quine's doubt concerned the very existence of a priori knowledge. If an *a priori* truth is one whose truth is necessary, an a priori statement can never be falsified. But Quine supported the view that "no statement is immune to revision." [7] It is arguable that reasonable revision is not always owing to error and that Quine's claim, if sound, does not necessarily undermine the possibility of genuine a priori knowledge. Yet Quine did seem opposed to the idea that genuine a priori knowledge is attainable. Although most philosophers nowadays seem to disagree strongly with Quine on this matter, he raised what I regard as the fundamental issue about a priori knowledge, and I shall pursue it later in the chapter.

Even if the existence of genuine a priori knowledge is not a serious issue for us, we must come to terms with the fact that Kant seems to have been wrong in holding that a priori knowledge is invariably universal and necessary. Saul Kripke made a strong case for this in lectures he gave in 1970 and later published as *Naming and Necessity.*[8] Some acute philosophers have raised objections with Kripke's criticism of Kant's contention,[9] but if Kripke's argument is reconstructed as follows, I think it is successful. Consider the assertion that the length in meters of a certain metal rod, the one known as the standard meter, = 1. Call this rod "*r*" and assume that we are speaking of it as it was at the time it was adopted as the official standard for measuring in meters. At this time, *r* had a particular length, call it "*L.*" According to the standard officially adopted, the length in meters of an object x at a time t is equal to 1 just in case x has L at that time. Expressed symbolically, this consequence of the standard is as follows:[10]

[SM]: For all x and t, $L_m(x,t) = 1$ if and only if x has L at t.

To show that the length in meters of the rod *r* is now, when the standard is adopted, equal to 1, we need only apply the rule SM to *r* itself. Instantiating the variables of SM to *r* and now, we obtain the consequence:

---

[5] *Ibid.*

[6] Quine (1953).

[7] Ibid, p. 43.

[8] See Kripke (1980), pp. 56, 122n. In the early 1970's David Kaplan pointed out that an utterance of "I am here now" is analytically true although it is not (or does not state) a necessary truth. See Kaplan (1992a), pp. 508ff. The analytic truth of this utterance depends crucially on the fact that the referent of "here" is not determined by something other than the utterance in which it occurs. As Frank Jackson observed, if I point to a place on a map when I say "I am here," I might say something false. See Jackson (2000), p. 332.

[9] See Soames (2004), ch. 16.

[10] A standard of measurement for an extensive quantity (of which a meter is an instance) requires a much more complicated convention than SM, but SM is sufficient for the argument at hand. For a very helpful discussion of what such a convention actually requires, see Carnap (1966), chs. 6-9.

[C] $L_m(r, \text{now}) = 1$ if and only if *r* has L now.

Since we have stipulated that the referent of "L" is the length *r* now has, we know that *r* now has this length. We may therefore infer from C that $L_m(r, \text{now}) = 1$ or, in English, that the length in meters of r is now = 1. Our knowledge of this conclusion is a priori because we obtained it from a stipulation identifying L and a standard for determining whether a thing's length in meters is or is not equal to 1.

Although we know a priori that this result is correct, what we know is not a necessary truth. It is not necessary that the length in meters of *r* is now 1 because the length of *r could have been different* from L at this time. If *r* had been heated, it would have a length longer than L; if it had been cooled in a significant way, it would have shorter length. Thus it is *possible* that *r* has and always had a length that differs from *L*. As things are, the length in meters of *r* is equal to 1 because the length *r* happens to have was arbitrarily chosen as the standard for measuring lengths in meters. If *r* had possessed a different length, the convention would have been different if that length had been adopted as the standard unit. But *r*'s length was not different and the standard was not changed. So we can have a priori knowledge of something that is actually contingent.

In his 1970 lectures Kripke also argued that Kant was wrong in thinking that necessity is a criterion of a priori truth—that if a truth is necessary it must be knowable a priori.[11] Consider the assertion, "The person who in fact discovered bifocals was Benjamin Franklin." It is possible that Franklin did not discover bifocals, but if we know that he was the one who in fact did discover them, we can use the description "The person who in fact discovered bifocals" to single him out in actual as well as possible or, as Kripke called them, counterfactual situations. Now Franklin was necessarily himself: *he* could not possibly have been someone else. If we know, then, that the person who in fact discovered bifocals was Franklin, then we know that this person, Franklin, was necessarily Franklin. We therefore know that, on this reading, the statement "The person who in fact discovered bifocals = Franklin" is necessarily true. But the necessary truth of this statement cannot be known a priori. To know that it is necessarily true we must know that the description "The person who in fact discovered bifocals" applies to Franklin, and this is a matter of fact that can be discovered only empirically.

If we set aside identity statements and contingent statements that, like "The standard meter is one meter long," can be known to be true merely on the basis of conventions about meaning, we can perhaps agree with Kant that a priori truths, if they exist, are universal and absolutely necessary. The question is, "How can we possibly know a priori that any statement of this kind is true?" How could we know such a thing at all? I noted that W.V.O. Quine seemed to believe that this kind of knowledge is not actually attainable. He may have been wrong about this, as most philosophers apparently now believe, but the question is certainly important. How is such knowledge possible?

According to tradition, a priori knowledge is either axiomatic or provable by necessary inferences from axiomatic premises.[12] The idea of being provable this way can be made more precise by the following definition:

> A proof for a proposition P is a finite sequence of formulas ending in P each of which is either an axiom or an elementary logical consequence of preceding formulas.

---

[11] Kripke, pp. 97-105.

[12] The classic account of this is given by Descartes in "Rules for the Direction of the Mind" (written in 1628 or thereabouts); see Descartes (1985), p. 14. In recent times Roderick Chisholm expounded a similar idea; see Chisholm (1996), ch. 3.

The formula ending the sequence here is a conclusion proved by the formulas preceding it. Since the sequence is finite, the formulas preceding the conclusion have an initial member. If we allow that the sequence may have only a single member, then P must be an axiom itself—in which case we can say that every axiom is a proof of itself. Also, if we allow conditional proof or indirect proof as elementary forms of valid inference, we can allow in proof-sequences formulas that are not inferred or even inferable from axioms. If these forms of inference are not counted as elementary, the conclusion of a strict proof will be inferred only from axioms and their logical consequences.

If the traditional idea is right, then, a priori knowledge will depend on or be obtained from axioms and elementary forms of inference. The requisite forms of inference must obviously be truth preserving: when they are applied to true premises, the conclusion they permit must invariably be true. If we can have a priori knowledge of something that is not itself an axiom, we must *know* that these forms of inference are truth-preserving. But how can we know this? The traditional answer is that the truth-preserving property of these forms of inference is knowable in the same basic way that axioms are knowable. For epistemological rationalists, axioms are known to be true by direct intuition or rational insight, and elementary forms of valid inference are known to be truth-preserving by the same kind of intuition or insight. For empiricists, the standard view is that logical axioms are analytically true, or true by virtue of meaning, and the truth-preserving property of elementary argument forms is insured by corresponding semantical rules.

Axioms and Primitive Rules of Inference

The rationalist idea that axioms are intuitively obvious does not accord with current logical practice. In fact, no particular formulas are now universally or even generally recognized as logical axioms. There are many different systems of classical logic, and although the theorems of standard systems are always the same, the axioms (if any are chosen) and the primitive rules of inference are often significantly different. As an example, Bertrand Russell and A. N. Whitehead listed five axioms for the system of propositional logic that they included in *Principia Mathematica,* namely:

1. $(p \vee p) \supset p$
2. $q \supset (p \supset q)$
3. $(p \vee q) \supset (q \vee p)$
4. $[p \vee (q \vee r)] \supset [(p \vee q) \vee r]$
5. $(q \supset r) \supset [(p \vee q) \supset (p \vee r)]$.

Paul Bernays soon proved that axiom (4) could be derived from the others and that it was therefore redundant, not needed as an axiom. But Russell's friend Jean Nicod offered a further simplification. Instead of taking "~" and "∨" as primitive connective symbols for the system, as Russell and Whitehead had done, Nicod suggested that "|" be used as the sole primitive connective, the formula "P|Q" having the sense of "not both P and Q." If this convention were adopted, Nicod showed, the whole system could be based on a single axiom with "P, P|(Q|R) so R" as the single primitive rule of inference. The axiom Nicod gave is "[p|(q|r)]|([t|(t|t)]|{(s|q) |[(p|s)|p|s)]})." I doubt that anyone would say this axiom is self-evident.[13]

I mentioned that Russell and Whitehead used "~" and "∨," translated "not" and "or," as primitive logical symbols for their system. This choice is significant be-

[13] The example is discussed in Kneale and Kneale (1963), p. 26.

cause the symbol for "if…then," namely "$\supset$," which plays a dominant role in their axioms, has a technical meaning that can be defined by means of "$\sim$" and "$\vee$," which are sufficiently close in meaning to the familiar "not" and "or" to be "taken as primitive," that is, used without a definition.[14] If one were merely told that "$p \supset q$" is to be understood as "if p then q," one would probably have great difficulty understanding why "$p \supset (q \supset p)$" and "$\sim p \supset (p \supset q)$" should have the status of logical truths and why "$p \supset (q \supset p)$" could reasonably be adopted as an axiom. The technical definition of "$p \supset q$" as "$\sim p \vee q$" makes it obvious why these formulas are logical truths: both are equivalent to "$(p \vee \sim p) \vee q$", which is a tautology.

Because the truths of classical propositional logic are distinguished by properties that can be characterized in formal terms—for instance, all theorems of this logic are truth-table tautologies—the class of such truths can be identified independently of axioms and rules of inference. The point in identifying axioms and primitive rules of inference for this system is to systematize the class of its truths—to identify a small class of truths from which the other truths can be inferred. Doing this makes a logical system useful for evaluating deductive inferences (they are valid if their conclusions can be derived from their premises by means of the axioms and rules chosen) and for ascertaining the logical truth of specific formulas: proofs can be constructed for logical truths. In the propositional logic the validity of inferences and the logical truth of specific formulas can be ascertained automatically by an algorithm (by truth tables), so the apparatus of axioms and rules is theoretically dispensable. But a comparable algorithm does not exist for the full system of predicate logic, so the apparatus of axioms and rules is vital there. Particular inferences can be shown to be valid, generally speaking,[15] only by means of a proof, and a formula can be shown to be logically true, generally, only in the same way.

My claim that the apparatus of axioms and rules is vital for the system of predicate logic actually requires an important qualification. Strictly speaking, axioms are not needed for a deductive system: rules are sufficient by themselves. Systems of "natural deduction" normally dispense with axioms.[16] To prove by deduction that some simple tautology is a logical truth, one can proceed by conditional proof or indirect proof. If we define "$p \supset q$" as "$\sim p \vee q$", we can show that "$p \vee \sim p$" is a logical theorem by a two-step inference. We first use conditional proof (C.P.) to prove "$p \supset p$":

1.	p	assumption.
2.	p	1, repetition
3.	$p \supset p$	1,2, C.P.

Then we use "$p \supset p$" to derive "$p \vee \sim p$":

4.	$\sim p \vee p$	3, definition.[17]
5.	$p \vee \sim p$	4, commutation

[14] The meaning of the symbol "$\vee$" actually differs from "or" in important ways. The symbol "$\vee$" must occur between formulas (or independent clauses), but "or" can meaningfully occur between noun phases, verb phrases, adjectival phrases, and adverbial phrases. The symbol "$\wedge$" differs from "and" in corresponding ways.

[15] I introduce the qualification "generally speaking" because some inferences in predicate logic and some formulas of that logic can be evaluated by an automatic procedure: for example, "$(\exists x)(\exists y)Fxy \supset (\exists x)(\exists y)Fxy$" is a truth-table tautology. The point is that no automatic procedure is available for all cases. This was proved by Alonzo Church (1936).

[16] See Montague and Kalish (1964).

[17] The definition "$(p \supset q) \equiv (\sim p \vee q)$" holds for all formulas; in the proof "q" is replaced by "$\sim p$."

The possibility of dispensing with axioms in logic is worth mentioning in a discussion of epistemological rationalism because it shows us that self-evident truths would not be needed in logic even if they were available.

In spite of what I have been saying in the past few pages, some rationalist philosophers will insist that certain specific formulas do express self-evident truths and that certain elementary inference-patterns are self-evidently truth-preserving. These formulas and inference patterns deserve to be accepted without inference, they say, and they deserve to be considered axioms and elementary valid argument forms whether they are actually treated this way by logicians, or not. The philosophers who argue this way usually support their case by citing certain examples—typically, the law of non-contradiction, the law of excluded middle, and the rule of *modus ponens*—but they never, to my knowledge, support their conviction that *all* logical truths can be derived from self-evident axioms and self-evidentially truth-preserving rules of inferences, nor do they explain how they could know such a thing. To nail down their rationalist position, the conviction must be rationally supported and the explanation must be given.

In all cases that I am aware of, the range of examples that rationalists cite to support their position is limited and narrow. Some of the examples are logical, some are mathematical, and some are metaphysical, "Nothing could be both red and green all over the same time" being a standard instance of the latter group. I shall argue that the instances they cite invariably lack any claim to self-evidence. In the next chapter I shall discuss some of them again, arguing that their truth—if they deserve to be considered true—can be supported by considerations favorable to empiricism.

General Doubts about Intuitive Knowledge

Earlier in this chapter I used the words "intuition" or "rational insight" to describe the kind of awareness rationalists claim to have of the truths they consider self-evident. These words are in fact very widely used at the present time, although their meaning is much less clear than their users suppose.[18] Before discussing the examples I mentioned in the last paragraph, I want to make some preliminary remarks about intuitive knowledge. The remarks are prompted by the fact that philosophers who speak of such knowledge apply the classification to some occurrences that empiricists have no trouble acknowledging. The examples that arouse empiricists' doubts and suspicions have distinctive features that are responsible for their negative attitudes. It is important to understand what these distinctive features are.

It is useful to begin with the cases that an empiricist would have no problem accepting. These cases include the recognition of particular things and the recognition of instances of kinds or qualities. As far as particular things are concerned, I can obviously recognize my face in a mirror, my wife in a crowd, or an old friend in a photograph. Normally, I recognize such things immediately; I do no inferring at all. Kant described these recognitional acts as intuitions, but their objects are not truths, and there is nothing dubious about them. The recognition of an instance of something is a little more complicated. When I recognize a color, I am recognizing that something I see, some particular thing, has that color; I do this when I see a flag to be blue and yellow. Here again the recognition is immediate. I recognize the instance (at least I often do so) without making any inference or drawing any conclusion.

Lawrence BonJour gives a slightly more complicated example; he describes it as an example of "rational insight":

---

[18] This is persuasively argued in Hintikka (1999).

> Even to apply as straightforward and seemingly unproblematic a rule as *modus ponens,* I must see or grasp in an immediate, not further reducible way that the three propositions comprising the premises and conclusion are of the right forms and are related in the right way….[19]

As an empiricist, I can easily grant that I *may* see or grasp in an immediate, not further reducible way that the three propositions have certain forms and collectively constitute an instance of *modus ponens*. Recognizing such a thing is something I have learned to do, and there is nothing philosophically problematic about this—nothing, at least, that I recognize as philosophically problematic. But I would emphatically deny that I *must* see the argument form in this immediate way. If the argument were composed in a language I read with difficulty, such as German or classical Greek, I would no doubt have to do some serious inferring to recognize the instance. I would probably have to look up a number of words and I might have to think about declensions, conjugations, or even genders before I could make the relevant identification.

BonJour's view on this last matter is reflected in a clause that he adds to the quotation above. His addition, which follows a colon, is this:

> that, for example, the two simpler propositions in question are in fact identical with the antecedent and consequent of the conditional proposition [*sic*] is as much a necessary, a priori knowable truth as anything else.

There appears to be some difficulty with the text here, for the initial "that" seems to be preceded by a tacit "I must see or grasp," which introduces the clause I cited in the last paragraph. But BonJour's thought, pretty clearly, is that the following is a necessary, a priori truth: "the two simpler propositions are in fact identical with the antecedent and consequent of the conditional proposition." As far as I can see, the truth of this assertion is certainly not knowable a priori. If anything is a necessary a priori truth here--apart, that is, from the conditional statement corresponding to *modus ponens*--it is only the conditional assertion, "If the argument is an instance of *modus ponens*, it consists of three statements, one a conditional and the others synonymous with the antecedent and the consequent of that conditional." And this is a general assertion, one that an empiricist would regard as analytic.

Other examples of recognizing things as such and such (recognizing x's as F) are generically similar to recognizing an instance of *modus ponens*. If we recognize that a certain sentence is or is not grammatical (in relation to our own language, dialect, or idiolect) or that the predicate "knows that P," as we understand it, is not applicable to someone who has no evidence that P,[20] our recognition may be immediate, but it is not philosophically troublesome. It results from a competence we have developed as we learned (or otherwise came to possess) the relevant verbal system. Forty years ago psychologists specializing in learning theory would have accounted for this competence by appealing to some "conditioning" process of stimulus and response; today, a favored explanation would no doubt advert to neural activity and innate verbal capacities. The phenomenon is straightforwardly empirical, and the best explanation will be empirical as well. Nothing here should raise the hackles of a responsible empiricist.

The examples that do raise problems concern alleged truths that are non-empirical. These supposed truths are problematic, empiricists say, either because it

---

[19] BonJour (1998), pp. 131f.

[20] Bealer (1999) regards this case as an example of a "rational intuition" (see p. 30).

is doubtful that anything genuinely factual is actually being recognized, or because it is far from clear, if something definitely factual and not merely verbal is being recognized, how that fact can possibly be known in the direct way rationalists suppose. Consider the first alternative, since I shall be discussing the second one in the next section. An example illustrating the problem empiricists see here can be drawn from the subject of ethics. Some philosophers nowadays attempt to prove the objective truth of certain moral judgments by pointing to examples that every reasonable person would acknowledge to be morally wrong—for instance, some young hoodlums setting a cat afire just for the fun of it.[21] But the moral judgment a reasonable person would make in a case like this and the repugnant attitude that would accompany it hardly show that an objective moral truth is being apprehended. Moral attitudes are instilled in children by parents and playmates as well as by pastors, teachers, and neighbors, and moral responses are evoked by these attitudes. Moral "facts" are poor candidates for the true explanatory factors, because different communities, and sometimes different groups in a larger community, instill different moral attitudes. On the day I write these words Islamic demonstrators in London are carrying signs declaring that the persons responsible for publishing a cartoon featuring the Prophet's face should be "beheaded." Yesterday a young man entered a gay bar in Boston and struck patrons with a knife and hatchet; and every day some people are demonstrating for, and others are demonstrating against, abortion, capital punishment, and the right to eat meat or use animals in medical experiments.

Apart from the variability of much moral opinion, people's opinions and feelings about what is laudable or blamable can be traced, as J. S. Mill emphasized in *On Liberty*, to "multifarious causes." One cause is, of course, the moral indoctrination they received as a child; this case is reinforced by what Mill called "the magical influence of custom, which is not only, as the proverb says, a second nature but is continually mistaken for the first."[22] Other causes that Mill cites include persons' reason or moral reflection, their prejudices, superstitions, and social affections, "not seldom their antisocial ones; their envy or jealousy, their arrogance or contemptuousness; but most commonly their desires or fear for themselves—their legitimate or illegitimate self-interest." Moreover, "wherever there is an ascendant class," Mill adds, a large portion of the morality of the country emanates from class interests and its feelings of class superiority"—and he supports this claim by a list of instances. These "multifarious causes," and others that Mill discusses, such as those occasioning the *odium theologicum* in a sincere bigot, which he takes to be one of the most unequivocal cases of moral feeling, make a rationalist's claim about directly apprehending moral facts seem decidedly simple-minded.

These remarks about the variability of moral judgment and feeling, and the many causes that bear upon them, do not, of course, imply that there is really no right and wrong in the moral domain. That is a contention that requires much further investigation. Mill himself seems to have believed that the utilitarian morality he accepted on purely secular grounds is rationally defensible and has a kind of objectivity,[23] but he did not suppose that the required defense included an episode of moral perception. Quite the contrary. If the defense he gave for the moral principle he advocated in *On Liberty* is representative of the defense he considered generally necessary, the requisite reasoning would be wide-ranging and elaborate. His argu-

---

[21] The example appears to be standard in recent discussion; I saw it in an unpublished paper by Ernest Sosa, and a similar example is cited in Hintikka (1999), p. 137.

[22] Mill (1859), p. 5.

[23] I am thinking here of his famous claim in *Utilitarianism* that "considerations may be presented capable of determining the intellect either to give or to withhold its assent to the doctrine, and this is equivalent to proof." See Mill (1861), p. 5.

ment in *On Liberty* is possibly the most complicated argument he ever developed for a single principle.

Another important source of doubt about the truth of what may seem intuitively obvious is the history of the axiom of parallels in Euclidian geometry. This axiom was commonly perceived to be less obvious than Euclid's other axioms, but some mathematicians believed they could derive it from them. When their derivations were examined by means of the more rigorous logical methods that became available in the latter half of the nineteenth century, critics discovered that the derivations made use of geometrical intuitions that were equivalent to the axiom of parallels itself.[24] These equivalent intuitions seemed so natural that they were not recognized as distinct principles. When, later in the century, mathematicians were able to prove that the axiom of parallels is in fact independent of the other axioms, systems of non-Euclidean geometry were worked out with different axioms in place of the parallels one.

In 1915 Einstein developed his general theory of relativity, according to which the geometry of physical space has a particular non-Euclidian structure. Well before this time Bertrand Russell had already distinguished "actual" space from mathematical space, holding that the study of actual space is "an experimental science."[25] As a branch of pure mathematics, he said, geometry is a subject whose assertions are to the effect that "such and such consequences follow from such and such premises, not that the entities such as the premises describe actually exist." Thus, he continued:

> If Euclid's axioms be called A, and P be any proposition implied by A, then, in the [old] geometry…, P itself would be asserted, since A is asserted. But nowadays the geometer would only assert that A implies P…. And he would have other sets of axioms, A_1, …, A_n implying P_1, …, P_n… respectively: the *implications* would belong to Geometry, but not A_1 or P_1 or any of the other axioms and propositions (pp. 373f).

A rationalist philosopher who can concede Russell's claim that the study of physical space belongs to empirical science might nevertheless argue that pure geometry is not essentially hypothetical but makes categorical assertions about ideal geometrical objects such as trangularity, squareness, and Euclidean parallelism. But this approach is no longer taken seriously by geometers. Geometry can do quite well without postulating such entities. Arguments for ideal objects are not mathematical arguments, anyway; and it is mathematically sufficient to hold that any thing or things satisfying the axioms of a given system, if there be such, must satisfy the theorems deducible from them. There is no need to go further than this.

The striking dubiousness of supposed intuitions in ethics and geometry should make a cautious philosopher highly suspicious of every appeal to intuitions. It is simply all too easy for people to convince themselves that they are in direct connection with the truth when they are merely imagining that they are so connected. But particular subjects may provide better candidates for intuitive knowledge than ethics and geometry. In the following sections I shall consider a representative sample of the examples rationalists now offer for what they consider directly self-evident. I begin with logical truths and primitive rules of a priori inference.

---

[24] See the clear and illuminating discussion in Carnap (1966), ch. 13.

[25] Russell (1902), p. p. 372.

Logical Truths and Rules of Inference

Perhaps the most frequently cited instance of a self-evident logical truth is the principle, or "law," of non-contradiction. Formulated in the usual way, "~(p ∧ ~p)," it seems to be a very simple principle, a suitable object of intuitive insight, but the formulation is very misleading.[26] The ingredient letter "p" is schematic; it stands in place of infinitely many formulas of infinitely varying complexity--and this infinite variety is a very inappropriate object of mental vision: we do not apprehend all the instances.[27] As a matter of fact, when we think about possible members of this infinite variety, some can be brought to mind that appear to falsify the law. Suppose we consider two statements, A and B, the first inscribed in a circle and the second in a rectangle. Suppose A is "The statement in the rectangle is true" and B is "The statement in the circle is not true." By obvious principles of logic and semantics we can easily derive the contradiction.

It may be useful to say a little more about how this contradiction is derivable. One way of proceeding is to use conditional proof. We first assume the hypothesis A and then derive its negation, from which we infer "~(The statement in the rectangle is true)." (From "A ⊃ ~A" we may infer "~A", for any formula "A".) We then assume the negation of A, that is, ~A, and proceed to derive A, from which we infer A, "The statement in the rectangle is true." (From "~A ⊃ A" we may infer "A".) We then conjoin the results of these inferences and obtain our contradiction, "The statement in the rectangle is true ∧ ~(The statement in the rectangle is true)."

As it happens, many sentences can be constructed in English that provide apparent counter-instances to the "law" of contradiction. Are they acceptable counter-instances? Obviously, someone convinced of the inviolable truth of the law of contradiction would want to say no, but to support this answer he or she will have to locate the error in the sort of reasoning I have given. When one uses a directly self-referential statement such as "This statement is false," one is apt to hear the response, "Self-referential statements are not acceptable substituends for the schematic letters in logical principles." But why should we accept this response? We can make all sorts of true statements that are directly referential, as when we say, "This is a sentence of English," "This is a grammatical sentence," and so on. Why are these sentences all right and sentences such as "The sentence in the rectangle is false" not all right when it appears in a certain circle and there is another sentence in a certain rectangle consisting of the words "The sentence in the circle is not true"? The only thing wrong is that these sentences or these combinations of sentences occurring in certain places give rise to contradictions.[28] If we want to avoid contradictions we can disallow such troublesome sentences, but we cannot plausibly rule them out on the basis of an alleged direct intuition of the truth of a law of contradiction. The alleged direct intuition is spurious because the law is highly general and schematic, and the intuition did not encompass (or survey) every instance pertinent to that law, some of which appear to provide demonstrable exceptions to it.

---

[26] Laurence BonJour, a recent defender of epistemological rationalism, formulates it with a quantifier in BonJour (1998), p. 33; he offers "for any proposition P, not both P and not P". But if his quantifier is understood in the usual way, his formulation does not make sense, for the inner formula then lacks a verb, like "not both Tom and Mary."

[27] This is the usual way of formulating the law; another way of formulating it is to take the axiom as expressing a particular proposition and to include a rule of substitution that permits one to obtain all the other instances I speak of in the text.

[28] Kurt Gödel's famous incompleteness theorem is founded on a formula that, in effect, says of itself that it is not provable in a system of a certain kind. Although the formula is self-referential, it is not considered paradoxical or objectionable because no contradiction is inferable from it. See van Heijenoort (1967), p. 352.

The reader should know that some respectable and responsible logicians contend that some statements should be accepted as both true and false because they can be proved to have this status.[29] Accepting them obviously requires someone wanting to retain a version of the principle of contradiction to restrict its application to formulas that cannot have both values. It is useless to object that this kind of restriction must be disallowed because any counter-instance to the classic principle will have to assume the principle in an unrestricted form. This objection is useless because it is false: asserting that "Q ∧ ~Q" is a counter instance to the schematic principle "~(p ∧ ~p)" does not involve assuming this principle. The classic principle, as commonly understood, is used to make an assertion about *all* conjunctions of a formula with its own negation. Asserting that a particular conjunction, "Q ∧ ~Q," is incompatible with "~(p ∧ ~p)" does not involve a general assertion of this kind.

Another standard logical principle often claimed to be intuitively obvious is the so-called law of excluded middle, "p ∨ ~p." Some rationalists actually doubt this principle, but it is inferable almost immediately from the principle of non-contradiction by one of De Morgan's laws. One reason for doubting the law of excluded middle lies in the vagueness of certain statements. A vague statement contains a predicate that clearly applies to some actual or imaginable objects, clearly fails to apply to other such objects, and neither clearly applies nor clearly fails to apply to a final group.[30] Objects in this last group are neither included in nor excluded from the extension of the predicate because the application conditions for the predicate are insufficiently definite to accommodate them. To take a proverbially vague predicate, suppose that Tom X is a man who is intermediate between being bald and being non-bald. Suppose that he cannot be truly classified either way. He is a borderline case of a bald man. If this is so, the semantic value of "Tom X is bald" is neither T (true) nor F (false) but IND (= indeterminate). But if "Tom X is bald" has this value, what is the value of "~(Tom X is bald)"? Obviously, it is IND as well. If it had the value T, "Tom X is bald" would have the value F; and if it has the value F, "Tom X is bald" would have the value T—and we are supposing that it has neither of these values. Well, if both "Tom X is bald" and "~(Tom X is bald)" have the value IND, what should be the value of their disjunction, "Tom is bald ∨ "~(Tom X is bald)"? The value should be IND as well. If neither of the subformulas has the value T, their disjunction can hardly have the value T: disjuncts with at least one true disjunct have this value. Similarly, it cannot have the value F, because formulas with false disjuncts have this value. But the law of excluded middle requires that it have the value T. Thus, the law appears to fail for this conjunction; it does not hold for all cases.

A similar result obvious holds for the principle of non-contradiction. If the value of both "Tom X is bald" and "~(Tom X is bald)" is IND, the value of their conjunction must be IND as well. The value could not be T because conjunctions with this value have true conjuncts; the value could not be F because neither conjunct has this value—and at least one must have it if the conjunction has it. But if the principle of non-contradiction is a law, the conjunction must have the value T. Any other value for the conjunction, for example IND, would be an objection to it. A person rejecting the law of contradiction as universally valid need not claim, therefore, that some instance of the schema "~(p ∧ ~ p)" is false and that an inner formula of

---

[29] See Priest (1998).

[30] See Sorenson (2006). Sorenson and Williams (2000) think vague statements should be considered true or false, but their reasons for thinking this have to do with the advisability of retaining classical logic. I comment on this matter in chapter three; see footnote 23.

the form "p ∧~ p" is true. It is sufficient to claim that some instance has a value other than T or F.

Philosophers who claim to see directly that the basic laws of classical logic are true obviously overlook ostensibly contrary instances involving vague predicates just as they overlook contrary instances containing "is true" and "is false." They do not see or intuitively apprehend the full generality of those laws; they do not contemplate all the instances pertinent to them, some of which appear to provide counter-instances. They neglect these instances just as they neglect currently rejected or questioned principles once deemed self-evident by other rationalist philosophers. Frege's Axiom of Abstraction[31]—that a class corresponds to every property—is now firmly rejected although it was once widely accepted as a truism; and the principle that every occurrence has a cause is now commonly regarded as false on scientific grounds although philosophers never seriously doubted it before the middle of the twentieth century. One would think that if highly respected philosophers had made erroneous claims about what is self-evident or intuitively obvious, their claims that this or that proposition has this status should be taken with a thousand grains of salt.

Do I believe that I have refuted the principles of classical logic? Do I think they should be rejected as false? It depends on how they are interpreted. If the principles—that is, the theorems—of classical logic are supposed to hold for all grammatical sentences that can possibly be put in place of the schematic letters in those theorems, which is the way many philosophers seem to regard them,[32] then the answer is yes. I have cited examples that will then count as counter-instances. But the principles of classical logic need not be taken to hold for all such sentences; they can be understood as applying to a restricted class of sentences, the ones that can be described as *proper substituends* for the schematic letters. If the principles are understood this way, it is arguable that I have identified no counter-instances. The crucial issue, then, is then how the proper substituends are identified. If we exclude the cases I have identified, have we excluded all possible falsifying instances? We certainly cannot claim to know this by "intuition," for we have not consciously surveyed all possible cases. It is uncertain, off-hand, whether statements such as "Zeus is insane" or "The Easter Bunny has a good sense of humor" conform to the principle of bivalence and are therefore either true or false but not some third value. This uncertainty is philosophically significant because, to be certain that we have rightly identified the class of sentences, or formulas, for which the theorems of classical logic are certain to hold true, we must know that no further qualifications will have to be made, and it is not at all obvious how we are suppose to know this. I will pursue this matter in the next chapter, when I consider an empiricist approach to logical truth. My aim in this chapter has been to cast serious doubt on the rationalists' approach. I think I have clearly said enough to make their strategy of directly intuiting the truth of a logical principle seem patently unrealistic.

To assure the reader that the opinion I am expressing here is not idiosyncratic, it is worth mentioning the example of Kurt Gödel, whose opinions on logic de-

---

[31] This is commonly formulated as an axiom schema, "(∃C)(∀x)(x ∈ C ≡ Φx)." Taking "Φx" as "x ∉ x", one can quickly derive a contradiction. See Suppes (1960), pp. 5-8.

[32] This seems to be true even of such well-informed and able philosophers as Hartry Field (2005), who very recently expressed the opinion that defenders of classical logic and of alternatives such as "fuzzy logic" do not really disagree as to whether any instances of excluded middle are true; the fuzzy logician will "just refrain from asserting some" (p. 84). Field is clearly wrong about this. At most the fuzzy logician will deny that he can point to an instance of excluded middle or non-contradiction that is actually false. He can, however, point to an instance that is plausibly *not true,* and this is enough to motivate his interest in an alternative to classical logic. If he is a firm believer in fuzzy logic, he will contend that what is plausibly not true in regard to vague statement is actually not true.

serve everyone's respect. In commenting on *Principia Mathematica*, which he acknowledged to be the "first comprehensive and thoroughgoing presentation of a mathematical logic and the derivation of Mathematics from it," he expressed his regret that the work "is so greatly lacking in formal precision in the foundations." What is missing there "above all," he said, "is a precise statement of the syntax of the formalism." He illustrated this lack of precision by pointing to Russell's treatment of what he, Russell, called incomplete symbols, such as definite descriptions. Russell introduced such symbols by rules describing how sentences containing them are to be translated into sentences not containing them. But, Gödel said, "to be sure … that (or for what expressions) this translation is possible and uniquely determined and that (or to what extent) the rules of inference apply also to the new kind of expressions, it is necessary to have a survey of all possible expressions [of the relevant language system], and this can be furnished only by syntactical considerations."[33] We need such a survey to be sure that even a rule such as non-contradiction or excluded middle applies to every sentence of the language we are using.

Before pushing on to a consideration of the non-logical examples I promised to discuss, I want to say something about two elementary rules of inference, *modus ponens* and *modus tollen*s. Rationalists typically regard these rules as self-evidently acceptable, but there are examples that some philosophers have considered counter-instances to them. I cite a possible counter-instance to *modus tollens* first, since it is the simplest:

> If it rained yesterday, it did not rain hard (yesterday).
> It did rain hard (yesterday).
> Therefore, it did not rain yesterday.

This seems to be a clear case of *modus tollens*, yet some have found it sufficiently problematic to merit discussion in a well-known philosophy journal.[34] An intuitive glimpse is evidently not sufficient to assure every sober mind of its indubitable status. Both premises could be true, but the conclusion must be false if the second one is true. A counter-instance to *modus tollens* might therefore seem to be a possibility.

Vann McGee discovered my second example some years ago.[35] Intuitively, it is much more plausible than the first example. To appreciate it, recall that the 1980 presidential election was won by Ronald Reagan, a Republican, and that Jimmy Carter, a Democrat, was second and Anderson, a Republican running as an Independent, was third. The example concerns this election:

> If a Republican wins, then if Reagan does not win, Anderson will win.
> A Republican wins (=does win).
> Therefore, if Reagan does not win, Anderson will win.

The first and second premises seem obviously true: Reagan won, and he and Anderson were the only Republicans running in the election. But the conclusion seems false. The real race was between Reagan and Carter; Anderson was far behind. So at the time of the election it would be false to say, "If Reagan does not win, Anderson will win."

---

[33] Godel (1951), p. 126.
[34] See Adams (1988) and Sinnot-Armstrong *et al* (1990).
[35] See McGee (1985).

There is actually some controversy about whether this argument is a genuine counterexample to *modus ponens*.[36] Its author, Vann Magee, thinks it is a genuine counterexample. At least one writer, Christopher Gauker, thinks it is not a counterexample to *modus ponens* but thinks it can be converted into a counterexample to *modus tollens* by switching lines (2) and (3) and negating them both.[37] I think (for reasons I shall mention in the next chapter) that, without some clarification of the English in which the argument is cast, it is impossible to say decisively whether it is or is not an acceptable counterexample. Here I shall merely note that the disagreement about this argument and the earlier one involving *modus tollens* supports my contention that the validity of these argument forms is not something that can plausibly be immediately grasped by an act of rational insight. As before, too many formulas are involved; too many considerations arise; too much cannot be decided without examining actual cases.

Alleged Self-Evident Factual Truths

The following are representative examples of nonlogical truths that rationalists claim to be self-evidently true; similar examples were included in a list supporting rationalism in a very recent discussion.[38]

1. A square is a rectangle.
2. Red is a color.
3. Everything red is extended.
4. Nothing can be both red and green all over.
5. *Taller than* is a transitive relation.
6. 7 + 5 = 12.

Off-hand, one would think that the first three examples are true by definition. My desk dictionary defines a square as an equilateral rectangle, and this implies that the sentence means "An equilateral rectangle is a rectangle," which satisfies Kant's famous definition of an analytic truth. The word "red" clearly refers to a certain color, and one would think that colors are by definition properties of spatially extended objects or quasi-objects such as rainbows.[39] Thus, only the last three examples would seem to be initially plausible cases of truths that might be immediately known by rational insight.

Rationalists view these sentences otherwise, of course. According to Roderick Chisholm, perhaps the best-known defender of epistemological rationalism in the last half of the twentieth century, the words in these sentences stand for "properties" and the sentences are true by virtue of essential relations between these properties. If, Chisholm said, we understand these sentences, we know what the relevant properties are; and if we bring them to mind, we can grasp the essential, unchanging relations between them. Our grasp of these relations shows us that the sentences must be true.[40]

Lawrence BonJour expounds a more complex view of how we grasp the truth of a priori propositions in his book, *In Defense of Pure Reason.*[41] His initial statement

---

[36] Bernard Katz has also argued that the argument does not provide a successful counter-instance to modus ponens, but his criticism is not the same as mine. See Katz (1999).
[37] See Gauker (1994) pp. 141f.
[38] See BonJour (2005), p. 100.
[39] Actually, the third example raises an issue concerning spatial points that I will discuss in the next chapter.
[40] Chisholm (1996), p. 27.
[41] BonJour (1998).

of this view is very similar to the one Chisholm offers, but his elaboration of it introduces complexities that Chisholm did not consider. He proceeds by discussing the example of "seeing" that nothing can be both red and green all over at the same time. Like Chisholm he begins by emphasizing that in understanding the proposition he "comprehends or grasps the property indicated by the word 'red' and also that indicated by the word 'green'," and that he has "adequate conceptions of redness and greenness." He also claims to understand "the relation of incompatibility or exclusion that is conveyed by the rest of the words in the verbal formulation of the proposition, together with the way in which this relation is predicated of the two properties by the syntax of the sentence." Given this understanding of the ingredients of the proposition, he says he is able to "see or grasp or apprehend in a seemingly direct and immediate way that the claim in question cannot fail to be true--that the natures of redness and greenness are such as to preclude their being jointly realized (p. 101)."

His elaboration of this initial statement occurs seven pages later in his book:

> It is in the natures of both redness and greenness to exclusively occupy the surface or area that instantiates them, so that once one of these qualities is in place, there is no room for the other; since there is no way for the two qualities to coexist in the same part of the surface or area, a red item can become green only if the green replaces the red" (p. 108).

BonJour's initial account was directly perceptual: redness and greenness are somehow presented to his consciousness, and he sees their incompatibility directly. But this second account is more discursive. Seeing one thing, redness, he realizes that it is by nature a certain sort of thing--an exclusive occupier, with respect to a certain class of properties (color properties), of a surface or area. Seeing another thing, greenness, he realizes that it has a similar nature: it too is an exclusive occupier, with respect to the same class of properties, of a surface or area. Since he sees that redness is different from greenness and knows that both properties belong to the excluder class, he then *concludes* that no surface or area can be both red and green at the same time. How he realizes that redness and greenness are exclusive occupiers in this way is not obvious on this model. But when he does realize this, he concludes that redness and greenness cannot occupy the same surface at the same time by a valid form of reasoning, one requiring the complex premise, "For any x, y, z, and t, if x and y are exclusive occupiers, with respect to a class of properties C, of a surface or area z at time t $\wedge$ x $\neq$ y, then $\sim$(x and y occupy z at t)."

It seems to me that the elaboration BonJour offers makes his account much more realistic than the one Chisholm presents. Yet I also think that the tacit inference required by the elaboration is best understood and justified by the kind of empiricist position that I shall expound in the next chapter. To carry on my criticism of epistemological rationalism, I shall not restrict my target to the more complicated account but shall proceed as if the question of which rationalist view is most plausible were still up in the air.

A distinctive weakness of both BonJour's and Chisholm's views of the examples I have listed is their undefended assumption that redness, greenness, and color are discrete properties that we can "grasp" in the immediate way they describe. There is a long tradition of thinking of color this way (G. E. Moore famously described yellow as a simple, unanalyzable property),[42] but the concept we use is much too complicated for such a picture. For one thing, a surface can be red but appear to have some other color if seen in some atypical light. To be the way it looks a red

[42] See Moore (1903), p. 10.

object must look red when viewed in good light by an observer with a good eye for colors. This fact about observers and conditions of illumination is built into the concept of red (or any objective color) and this makes the property of being red a very complicated one. Because of this complexity, it is extremely doubtful that the connection of this property to a logically distinct property of comparable complexity could be grasped in any immediate, infallible way.

Another fact to keep in mind here is that the ideas (or concepts) expressed by the words "red," "green," and "color" are far too vague and too generic to represent discrete, graspable essences. If you start with a pail of white paint and begin adding small amounts of red, the paint will gradually become a faint pink, then darker and darker pink, and finally, if you add enough red, the paint will start getting red and eventually be red. There will be no cut-off points indicating when the paint first becomes pink and then stops being pink, and when it first becomes red. The same is true of red and many other colors that can be blended with it: there are no natural cut-off points that define the compound shades that may result. If a *de re* correlate were needed for "red," the most plausible candidate would seem be a so-called fuzzy set whose positive members include numerous shades of red (no doubt many thousands of them) of various degrees of brightness and saturation blended with wide variety of other colors.[43] The set would be fuzzy because its membership is not categorically defined: things belong to it in greater or lesser degrees.

If things are definitely red only because of determinate shades that fall within a certain range, the property of being definitely red is a derivative one specifiable by a quantified formula such as this: "$\forall x$(x is definitely red iff $\exists P$(P belongs to the family of definitely red shades $\wedge$ x has P))." A property so specifiable is plainly not a plausible object of direct apprehension. It can be "grasped" only by a discursive process seriously at odds with the picture presented by Chisholm and BonJour.

In spite of the complexities I have been emphasizing, there is nevertheless a kind of incompatibility between red and green that has nothing to do with metaphysics. Owing to the physics of light and the physiology of the human eye, it is not possible for us to perceive shades that contain mixtures of these colors.[44] To expose the error in the rationalist's metaphysical claim about the incompatibility of different colors, it will therefore be useful to consider a different pair. Yellow and green will suffice. A detailed example will also be useful, because the subject in question is complicated in ways that I have yet to indicate.

Before I present my example I want to announce upfront the strategy I shall be pursuing. I aim to show, first, that there is really no plausibility in the idea that a surface or surface part could not jointly exemplify, all over, two different generic colors. Such things could conceivably be both generic green and generic yellow at the same time. I will concede that they could not equally possess two distinct shades of color, any color, at the same place at the same time. But this last impossibility is not a synthetic truth that is known in the intuitive way BonJour describes. It is rather, I will argue, an analytic truth that follows from (and is provable by reference to) a basic classificatory convention for identifying determinate color shades.

Here is the example. Suppose two people, Tom and Mary, visit an arboretum and see a shrub with leaves whose color appears to include both these colors. Tom and Mary are told that the color is chartreuse, but it does not satisfy the definition of that color given by their dictionary, which is "a clear light green with a yellowish tinge."[45] Tom describes the color as greenish-yellow, which is a shade of yellow; and

---

[43]For an informal discussion of fuzzy set theory and arguments for the view that the semantics of basic color terms is best represented in the formalism of fuzzy set theory, see Kay and McDaniel (1997).
[44]Ibid.
[45] This is what is given in *The Random House Dictionary of the English Language* (1968).

Mary describes it as yellowish-green, which is a shade of green. Considering the novelty of the color, neither person is clearly right, but their classifications seem inconsistent: greens and yellows are generically different. If both persons hold stubbornly to their own classifications, it seems reasonable to say that they are demarcating yellows and greens in different ways and that they therefore mean slightly different things when they speak of these two colors.

There is, however, another way of thinking about the color of the shrub that is no less acceptable than the ones Tom and Mary have. Tom, who describes the leaves as greenish-yellow, sees a kind of yellowness all over a given leaf; Mary, who describes them as yellowish-green, sees a kind of greenness there. But suppose a friend, Harry, describes the leaf as green and yellow all over: he sees both greenness and yellowness there. For him, the two colors are both present in this instance, and neither predominates. Instead of describing the color he sees as greenish-yellow, which is a shade of yellow, or yellowish-green, which is a shade of green, he describes it as green-yellow, a shade that exemplifies both generic colors in an equal degree. His descriptions shows that he conceives of generic green and generic yellow as overlapping in a region of the spectrum, and his conception makes it consistent for him to say that a thing can exemplify both colors all over at the same time.

As I said before I presented this example, I am not using it to provide a counterexample to a plausible rationalist claim about color-incompatibility. If two colors are described purely generically, there is really no plausibility in the idea that they cannot be exemplified by the same surface and the same time. Thinking this is impossible can only be owing to carelessness. But two specific color shades, no matter what generic colors they involve, are incompatible in the sense in question. No surface could possess them both at the same time. This fact does not support rationalism, however. It is simply a logical consequence of the way we distinguish specific color shades.

To appreciate this last fact, we should observe that a physical thing could clearly possess the same specific, absolutely determinate color at different times. But what would determine—what would settle the question—whether the absolutely determinate color (the shade that does not include more specific shades) possessed by a surface at time t is different from the determinate color it possesses later, at a time t*? The answer, pretty clearly, is "The determinate color the surface possesses at t is the same as the determinate color it possess at t* when and only when the color of the surface at the one time is indistinguishable from its color at the other time." But if two determinate colors are conceded to be distinguishable, it *follows logically* that nothing possesses both of them at the same place at the same time.[46]

It is important to realize that the impossibility at issue here is not a mere matter of color-exclusion; it is something that attaches to the truth of a conjunctive proposition, one that is best expressed by a sentence such as "There is an F, G, x, and t such that F is an absolutely determinate color, G is an absolutely determinate color, $F \neq G$, x is a part or the whole of the surface of a physical body, x has F at t, and x has G at t." I think there is no plausibility in the idea that the necessary falsity of this complex proposition could be known by the simple procedures that Chisholm and BonJour describe—this is, by bringing to mind and comparing different colors or color-shades. This is implausible because the complex proposition in question in-

---

[46]This is very easily proved. Formulate the principle of color identity for determinate colors (CIDC) as "$\forall A \forall B(DC(A) \wedge DC(B) \wedge \text{Distinguishable}(A,B) \supset A \neq B)$" and define "the determinate color of x at t" (in symbols, "DCxt = A") by "A is a determinate color and x has A at t." The impossibility of DC(xt) = A & DC(xt) = B & Distinguishable(A,B) (i.e. of x having both distinguishable A and B at a time t) follows almost immediately by conditional proof.

volves a technical concept—the concept of an absolutely determinate color—that presupposes distinctions that are not immediately obvious.

You might think that you could get an adequate idea of color determinacy merely by contemplating a color-expanse, since every actual expanse is bound to be fully determinate. But the sameness, the identity, of a determinate color is not a perceptible matter. An expanse of color may seem entirely homogeneous—I may be unable to discern any differences in it—but if one side is arranged against another, as when a side of a colored sheet of paper is folded upon another, I may learn that the color is not the same all over. Apprehending sameness always involves a comparison, so the identity of the color one sees (if there is just one) is not something that is immediately grasped by the mind. The notion of an absolutely determinate color, the sort of thing that cannot coexist with other properties of the same family, makes sense only in relation to a standard for sameness. The fact that the standard is objective (or public) indiscernibility in this case is not something that can be read off from what is before one's mind when one thinks of colors.

As I see it, the conjunctive proposition affirming the relevant incompatibility between colors is a consequence of a high-level analytic truth about absolutely determinate properties: it is a consequence of what we mean when speak of such things. I cannot properly defend this belief until I complete my discussion of analyticity, which I undertake in the next chapter. What I can confidently assert right now is that the BonJour and Chisholm account of how color incompatibilities are known is not credible. It is just as dubious as their account of logical truth.

Not all rationalists would agree with the accounts BonJour and Chisholm offer for the means by which color incompatibilities are ultimately known. George Bealer, who is an acknowledged rationalist, approach these incompatibilities differently, contending that intuitions of a priori certainty are episodes of "seeing" that are *prima facie* rather than certainly true.[47] But seeming incompatibilities are just as incapable of doing justice to the color-incompatibilties I have been discussing as the sort of direct perception of property connections that BonJour and Chisholm describe. The phenomena are bound up with conventions about the sameness of determinate colors, and they require a quite different analysis.

Although I shall pursue the notion of analyticity only later, it is pertinent to mention here that when Roderick Chisholm, in his classic textbook, attempted to refute the empiricist contention that assertions such as "Everything square is a rectangle" and "Being red excludes being blue" are analytic, he relied on Kant's eighteenth-century definition of an analytic truth.[48] But this definition was far out of date when Chisholm offered his refutation. In fact, Gottlob Frege explicitly called attention to the inadequacy of Kant's definition more than a hundred years before Chisholm's third edition was published, and leading empiricists left Frege's improved conception well behind in the 1930's.[49] Thus, although Chisholm's examples are not analytic in Kant's sense, it does not follow that they are not analytic in an improved sense that is more generally applicable.

Three Final Examples, Two Old and One New

The last two sentences on the list of alleged self-evident factual truths given at the beginning of this section concern the transitivity of the relation *taller than* and the

[47] Se Bealer (1999b), p. 247/
[48] Chisholm (1996), pp, 34-36.
[49] I discuss this in the following chapter.

identity of 5 + 7 and 12. My claims about the vagueness of "red" and "green" are also applicable to the example concerning "taller than." This last predicate is not nearly as transparent in meaning as one might initially suppose. There is no doubt that Wilt Chamberlain is taller than Yogi Berra and that a dwarf is not taller than a giant, but there are many pairs of objects for which the question "Is A taller than B?" has no more definite an answer than "Is Tom bald?" Consider this: Can a frog be taller than a tadpole or a wristwatch taller than a ring? Can a mountain be taller than a hill? Frogs, tadpoles, wristwatches, and rings have vertical dimensions, but they are not described as tall or short, and it is not clear that one can be taller than another. As for mountains, they can be tall but not short, and hills can be high or low. Can mountains and hills be compared for tallness? There is no definite answer to this. One can measure the heights of a mountain and a hill and declare that the one with the greatest height is taller than the other, but this way of speaking is not standard, and not clearly right or clearly wrong. The permissible arguments in the schema "x is taller than y" are not sharply demarcated. Yet if *taller than* were a discrete, determinate property that can be taken in by an intuitive act of consciousness, it should either be possessed by an ordered pair of objects or not possessed by it. We should not have any undetermined cases.

Even though "taller than" is a surprisingly vague predicate, it can be defined by other predicates, some comparably vague, in a way that shows the transitivity of *taller than* to be a consequence of a more basic transitivity, one involving the mathematical concept of *greater than*. If *a* is in fact taller than *b*, then *a* has a height that is measurably greater than the height of *b*. Let "*h(a*)" abbreviate "the height of *a"* and let "*_T_*" abbreviate "_is taller than_." If we define "xTy" as "h(x) is greater than h(y)"—in symbols, "h(x) > h(y)"—then we can prove the transitivity of *taller than* by proving that for any x, y, and z, if h(x) > h(y) and h(y) > h(z) then h(x) > h(z)." But the latter is a mathematical truth, one that is independent of the supposed graspability of the property *taller than.* The question I raised above about the indeterminacy of the permissible arguments in the formula "_ is taller than _" does not affect this proof of transitivity, because the formula expressing the transitivity of the taller-than relation is hypothetical: if "_T_" holds for the arguments of the antecedent, it holds for the arguments of the consequent. There is no need to worry about indeterminate cases.

The examples concerning *red, green, color*, and possibly even *taller than* have been seized upon by rationalists because they judge them to be necessarily true and to involve concepts that are not definable by means of simpler ones. If the predicates they feature cannot be defined this way, rationalists suppose, the specimen statements in which they appear cannot be "true by definition" and thus analytic; the empiricists must therefore be wrong about the basis for their truth. Yet our inability to define certain predicates, or find necessary and sufficient conditions for their application to suitable objects, need not be taken as evidence that such predicates connote indefinable properties, or any properties at all. Their application to objects may be justified by reference to properties--as when one justifies the application of "bald" to a man because he may have the property of being utterly hairless. But the property that does the justifying need not, as here, be the property supposedly expressed by the predicate. No such property may exist. In the course of learning English we come to apply "bald" to bare scalps, to deny it of hirsute ones, and to apply it to intermediate cases only with modifiers such as "slightly," "nearly," "almost," and "kind of." The sentence "If a man has no hair growing on his scalp, he is bald" may be necessarily true, but its truth does not depend on a property of baldness that an attentive mind can grasp and compare with other properties. It depends on the way we use the words involved and the instances we recognize as unqualified instances of bald and non-bald persons.

The final example, the one concerning the identity of 5 + 7 and 12, deserves a far more elaborate treatment than rationalists typically offer. There have been importantly different theories of mathematical truth, and according to possibly the leading theory since the time of Frege, mathematical truths are reducible to truths of logic and set theory. Empiricists sometimes say that the truths of set theory are basically the same as those of logic,[50] but even if the former are acknowledged to possess a distinct subject matter, the axioms on which they rest are not declared to be intuitively obvious. Quine discusses five different axiom systems for set theory, and he spends many pages discussing their relative advantages and disadvantages. At no point does he attempt to justify an axiom by appealing to its self-evidence.[51] It may not be possible, in the end, to interpret mathematical truths in accordance with alternatives that philosophers have historically debated, but an interpretation that attributed their verification to a perception of intuitive obviousness would certainly not accord with the cautious attitude of serious writers on the subject.[52] As far as I can see, the obvious truth of "5 + 7 = 12" adds no significant support to the rationalist thesis regarding a priori truth.[53] To provide such support, those arithmetical examples must be accompanied by a credible philosophy of mathematics.[54]

Some interesting examples not offered by Chisholm or BonJour but apparently contrary to empiricist doctrine were given by Kripke in *Naming and Necessity*. Kripke did not offer these examples as anti-empiricist, but they are naturally viewed that way. Colin McGinn says that one of them (and I am confident that he would say the same of the others) is inconsistent with the empiricist view that necessary truths[55] are invariably analytic and depend for their truth on the analysis of the words involved in them. The example McGinn mentioned concerned the necessity of his being born to a particular pair of biological parents. His biological parents were Joe and June McGinn, and if Kripke was right, he, Colin McGinn, *could not* have been born to anyone except Joe and June. But the necessity here, expressed by the words "could not," is "not a matter of the meaning of the name 'Colin McGinn'"—nor is it a consequence of the meaning of the remaining words in the sentence, that is, of "could not have been born to anyone except Joe and June."[56] Kripke's other examples concern the matter from which an object such as a chair originated and the substance of which it is made. He expressed one example in the words, "If a material object has its origin from a certain hunk of matter, it could not have had its origin in any other matter."[57] His other example is to the effect that if a table was originally made of wood, it could not originally have been made of another substance such as ice.

Kripke mentioned these examples in two different footnotes, and he did not therefore discuss them thoroughly. It is clear, however, as one can infer from his second example, that Kripke intended all three to have the form of conditionals. The example McGinn mentioned would thus be fully expressed as follows:

---

[50] I can distinctly remember Carnap saying, in a logic seminar I took from him many years ago, that Zermelo's *Aussonderung Axiom* "looked like a logical axiom" to him.

[51] See Quine (1969), ch. Xiv.

[52] See Suppes (1960), Introduction.

[53] "i = the square root of minus 1" is just as much a mathematical truth as "5 + 7 = 12," but i has always been considered an "imaginary" number. A philosopher's intuition seems a poor basis for distinguishing i and 5. i is considered imaginary because, according to the axioms for "real" numbers, a number such as i cannot exist.

[54] A very suggestive novel approach to mathematical truth, one that does not support rationalism, is outlined in Fine (2005).

[55] Contemporary empiricists will of course now concede that identity statements may be necessary but not analytic. See above, p. 40.

[56] McGinn (2002), pp. 96f.

[57] Kripke (1980), pp. 114f.

If the biological parents of Colin McGinn were Joe and June, then
Colin could not have been born to anyone other than Joe and June.

Although this sentence contains three proper names, its status as a necessary truth is a consequence of a more general principle that is expressible without any proper names—specifically:

∀x∀y∀z(the biological parents of x are y and z ⊃ x could not have been born to anyone other than y and z).

This last sentence, since it is wholly general in form, is not such an obviously unlikely candidate for the status of analytic truth as the one above. If it could be shown to be analytic, the one above could then inherit it, as "Aune is not both wise and stupid" inherits it from "No one is both wise and stupid."

Is there any plausibility in the idea that the general principle can be shown to be analytic? I think there is, at least if that principle is true and can actually be proved to be so. Kripke himself says that the related example that he stated fully, the one about the matter from which an object originated, is "susceptible of something like a proof," and if the proof he seemed to have in mind is sound, it is arguable that the result is an analytic truth. But to argue this, an acceptable conception of analytic truth must be developed. I shall therefore return to the example in the next chapter.

An Indirect Argument for Rationalism

Some of principal arguments supporting rationalism are indirect: they are intended to support rationalism by undermining its most widely held alternative. Arguments of this kind are reasonable if there is a strong probability or significant rational assurance that one or the other alternatives being considered is true or approximately true. Lacking this assurance, we must view these negative arguments as essentially motivational, advanced in the hope that they will induce your opponents to abandon their view in favor of yours. But for anyone who believes that some truths are universal, necessary, and knowable a priori, the only acceptable alternative to rationalism is some form of empiricism. So an argument against empiricism is very important for epistemology.

Empiricists agree that there are a priori truths, but they say that such truths are analytic, true solely by virtue of what is contained in a concept (Kant) or, roughly speaking, what is meant by ingredient words (Carnap). R. M. Chisholm opposed these views, insisting, in effect, that the statements empiricists consider analytic are not about ideas or words but about the world that ideas or words represent.[58] The world thus represented is what makes the statements true, he said: they are true because of what the world is like. As an illustration, consider the statement that Kant used in introducing the notion of an analytic judgment, "All Bodies are extended." This statement is true; it is so, Chisholm said, because *all bodies are extended.* If they were not extended, the statement would be false. Reality provides the relevant truth condition--not words or concepts, as empiricists suppose.

If the argument just given were a good one, it would have to apply to all statements or all judgments. But many statements are hypothetical. Formulated in modern notation, Kant's statement that all bodies are extended would be "∀x(x is a body ⊃ x is extended)." This statement would be true if no bodies happened to exist

[58] I say "in effect" because the argument Chisholm actually gives is directed against what he calls "linguisticism," the view that a priori statements are "essentially linguistic," true by virtue of what words mean or how they are used. But Chisholm would certainly have modified the argument to apply to a conceptualist view, one implying that a priori statements are true solely because of what is contained in certain concepts or ideas.

at all. In view of this, how could it be things in the world, what exists at a time, that makes this statement true? If Kant was right, the statement is true because the predicate is contained in the concept of the subject: the subject concept specifies the conditions a referent would have to satisfy, and those conditions include the conditions required for satisfying the predicate. Because of this, we are assured that if anything were to satisfy the subject concept, it would satisfy the predicate. The fact that makes the judgment true is not, in this case, something about the extra-conceptual world; it concerns a connection between subject and predicate. It is a purely conceptual thing.

Chisholm, who said it is "properties" that makes a priori statements true, used a different kind of example in his truth argument; he did not use the example I borrowed from Kant. His example was:

> The English sentence "Being square excludes being round" is true
> if, and only if, being square excludes being round.

Here the truth of the mentioned sentence does apparently depend on something that somehow exists—namely, the properties of being square and of being round.

In spite of its initial plausibility, Chisholm's argument seems to break down on examination, and it certainly does not apply to the full range of truths that empiricists regard as analytic. The first step in examining it is to ask, "What does it mean to say that being square excludes being round?" "What is the property of being round excluded from?" If Chisholm is right, the world contains both properties, so the presence of one of them in the world does not (or should not) exclude the other from the world. Obviously, what being square excludes being round from is any object containing squareness; more exactly, it excludes roundness from any place where at any time some object contains squareness.[59] Since this is what the exclusion obviously amounts to, the statement "Being square excludes being round" really asserts something about all objects and all times, namely:

> $\forall x \forall p \forall t$ (s round at place p at time t $\supset$ ~(x is square at place p at time t)).

But this statement is hypothetical, and its truth does not obviously depend on what exists in the world. An empiricist would say it is dependent on an incompatibility between what is contained in the concept of being round and the concept of being square.

Another negative argument against empiricism applies primarily to modern versions holding that analytic truths are statements that are true by virtue of meaning, statements being true sentences with fixed interpretations. Rationalists such as BonJour insist that one can grasp a priori truths that one cannot express in language, and others claim that even if every truth were necessarily expressible in some language or other, one *may* nevertheless grasp a truth that is not in fact expressed in language. If a priori truths were invariably true statements, and if some of the truths that are supposed to be grasped are a priori, these claims must be false. Are they? What reasons can rationalists offer in their support?

In a passage in his book, *In Defense of Pure Reason*, BonJour wrote of the darkish blue of two books on his desk. He does not have specific names for these blues, he said, and he has no other means of representing them linguistically, but he nevertheless knows directly that nothing could have both of them all over at the same time. What is before his mind when he knows this cannot be linguistic, because the blues are not linguistically represented (pp. 57f). The trouble with this argument is that BonJour actually expresses the crucial proposition in language and

[59] A complex diagram could contain roundness (or circles) and squareness (or squares) at several places

does so in way that is as adequate for him as "This pen is mostly white" is adequate for me now. His reader does not know the referent of his "these colors" any more than my reader knows the referent my "this pen," but each of us knows what the referents of his own words are and each of us understands the sentences in which he has included those words.

In another passage BonJour quotes with approval A. C. Ewing's claim that a person who is capable of forming visual images might well see the truth of propositions such as the one concerning green and red without having to put them into words. To accomplish such a feat the person would no doubt have to have the concept of incompatibility that BonJour speaks of elsewhere,[60] but reflection shows that further concepts, or ideas, are necessary as well, for the proposition involves time, space, universality, thinghood, predication, and modality. To grasp the alleged truth about the red and green, one must be capable of thinking the thought *No thing could be both red and green all over at the same time.*" I suppose it is conceivable (at least if Wittgenstein was wrong about private languages) that someone could think such a thought without having a conventional language such as English or French, but it is hard to see how we could entertain all the propositions we are supposed to be capable of entertaining if we did not have a system of concepts or ideas that corresponds to words, particles, and grammatical constructions of conventional languages.

This response to the Ewing argument does not vindicate a language-centered account of analytic truth, but it does vindicate the sort of idea-centered or concept-centered approach of older empiricists. I say "approach" here because the details of their theories may be erroneous or inadequate for the full range of truths that a contemporary empiricist would want to consider analytic. But since ideas or concepts must be acknowledged as having contents that can be shared, wholly or partially, with other ideas or concepts, Ewing's argument does not itself refute the kind of account offered by older empiricists. I will consider its application to a more up-to-date account at the end of the next chapter.

[60] See p. 47 above.

Chapter 3

EMPIRICISM AND THE A PRIORI

My aim in this chapter is develop an empiricist account of *a priori* knowledge and to defend it against objections raised by rationalists and anti-rationalist critics such as W. V. O. Quine. Since empiricists famously regard *a priori* truth as analytic, I shall offer here a clarification and defense of analytic truth. The clarification is needed because the upshot of Quine's influential criticism was that, for all its apparent reasonableness, a distinction between analytic and synthetic statements has not yet been satisfactorily drawn. The idea that such a distinction can be satisfactorily drawn is, he once said, a "metaphysical article of faith." My aim here is not only to draw a satisfactory distinction but to show that the notion of analytic truth, suitably clarified, provides a reasonable explanation of how a priori truths can have the universality and necessity that they are traditionally supposed to have.

Quine's Critique of Analytic Truth

In "Two Dogmas of Empiricism" Quine criticized three post-Kantian definitions of analytic truth.[1] The first one he criticized was essentially Frege's, though he did not identify it as such. His criticism was focused on the class of supposed analytic truths that, like "No bachelor is married," are not logically true. According to Frege, statements of this kind are analytic just when they can be proved to be true by general logical laws and definitions. Quine described these statements a little differently, saying that they can be turned into logical truths by "putting synonyms for synonyms," the synonyms being expressions (words, phrases) appearing in the *definiens* and *definiendum* of the relevant definitions. If the definitions are acceptable, these expressions must be "cognitively synonymous": with the exception of poetic quality and psychological associations, their meaning must be the same. But how, Quine asked, can the synonymy of two words be known in a particular case? Can this be known if the word "analytic" is not understood already? He argues that the answer is no, and proceeds to look about for an alternative definition.

Why did Quine think that the notion of synonymy could be understood only if the word "analytic" is understood already? His reasoning was this. The definitions needed for the demonstrations Frege described served as principles of substitution. If the predicate "is a prince" is defined as "is a royal son," then we may substitute the latter for the former in the logical truth, "A prince is a prince," and obtain another truth, which can be considered analytic—namely, "A prince is a royal son." Since the words that good definitions allow us to substitute for one another must be cognitively synonymous, a promising way of defining cognitive synonymy is by means of substitutions that preserve truth: If substituting W1 for W2 in any true statement containing W1 always results in another true statement, the words W1 and W2 must be synonymous: they do what a good definition permits. This strategy seems promising until one realizes that the full range of statements containing a word W1 will include statements that also contain the word "analytic" (for instance, "It is analytic that princes are royal sons") or words that, if empiricists are right, can be understood only by means of "analytic"—for instance, "it is necessary that." If any of these statements were excluded from the substitution test, the test would not identify synonyms. If they are allowed, we can apply the test only if we already understand what we are trying to understand or make sense of.

---

[1] Quine (1953).

Although the intimate connection between "being synonymous" and "being analytic" makes it inadvisable to try to define "analytic" by means of "synonymy," Quine's strategy in making sense of analyticity was nevertheless highly peculiar from the beginning. He initially noted that the statements held to be analytic "by general philosophical acclaim" fall into two classes, the first including logical truths such as "No unmarried man is married." He expressed no difficulty in understanding what a truth of this first kind is. "If," he said, "we suppose a prior inventory of *logical* particles, comprising 'no', 'un-', 'not,' 'if', 'then', 'and', etc., then in general a logical truth is a statement which is true and remains true under all reinterpretations of its components other than the logical particles."[2] But even this initial, limited clarification is peculiar in a discussion of what an analytic truth is. Kant's definition[3] was intended to show us why analytic judgments are true, but Quine's characterization of a logical truth assumes that we can recognize the truth and the resultant truth of statements that are true and remain true under all reinterpretation of their components other than the logical particles. This gives us no insight into how we know that the relevant statements are true.

The same holds for Quine's proffered account of the second kind of presumed analytic truths, the kind containing "All bachelors are unmarried," and his suggested strategy for defining synonymy. His suggestion was that analytic truths of the second kind are statements that can be turned into logical truths by putting synonyms for synonyms. This could work only if we had some independent means of recognizing logical truths. His strategy for identifying synonymous expressions had a similar limitation. We were supposed to consider whether the result of substituting one expression for the other in all true statements would be a true statement. But if we were wondering whether a candidate analytic statement "All princes are royal sons" is true, the question whether "prince" and "royal son" are synonymous would oblige us to consider whether the result of substituting "prince" for the first occurrence of "royal son" in "All royal sons are royal sons" is true—which is to say whether "All princes are royal sons" is true." The strategy would simply take us in a circle and get us nowhere.

A satisfactory definition of "analytic" should give us an understanding of why all analytic statements are true, the first kind as well as the second kind. Kant's definition did not apply to the class of logical truths, and it worked only for a small part of the other class.[4] The problem is to find a definition that works for the totality of both classes and also provides the understanding that an empiricist, an opponent of epistemological rationalism, desires. Quine considered two further definitions, or groups of them, but neither, as he understood them, appeared to work for all cases or provide the desired understanding. One definition (one member of the class he considered) was applicable primarily to artificial, formal languages, the idea being that a statement of such a language is analytic if its truth is a consequence of the semantical rules laid down for that language. The other definition was based on the notion of empirical confirmation, although Quine relates it to the Verification Theory

---

[2] Ibid, p. 22.

[3] According to the definition Kant gave in his *Critique of Pure Reason* (A6, B10), an affirmative judgment is analytic just when its predicate is contained, perhaps only covertly, in its subject concept. To ascertain the truth of such a judgment, one has only to become conscious, he says, of what is contained in the subject concept. If the predicate concept is affirmatively contained in the subject concept, the judgment must be true, because anything to which the subject applies will satisfy or fall under the predicate: the predicate will apply to it, too.

[4] As early as 1884, Frege emphasized that Kant's definition does not include relational judgments such as "If the relation of every member of a series to its successor is one- or many-one, and if m and y follow in that series after x, then either y comes in that series before m, or it coincides with m, or it follows after m." See Frege (1950), p. 103.

of Meaning: An analytic statement is one that is "confirmed no matter what."[5] This last definition is not credible in view of current conceptions of empirical confirmation,[6] but the "semantical rules" approach is far better than Quine supposed, and I will discuss it further in a later section of this chapter. Quine took a more moderate approach to analyticity in a later paper, and it will be instructive to consider his view in this paper next.

Quine's Later View of Analyticity

Forty years after he published "Two Dogmas..," Quine published **"**Two Dogmas in Retrospect."[7] In this later paper he summarized the more generous attitude toward analyticity that he had expressed in some of his later work. According to this more generous attitude, "analyticity undeniably has a place at a common-sense level… It is intelligible and often useful in discussions," he said, "to point out that some disagreement is purely a matter of words rather than of fact." A paraphrase that avoids a troublesome word can often resolve the disagreement. Also, in talking with a foreigner we can sometimes recognize "some *impasse* as due to his having mislearned an English word rather than to his having a bizarre view of the subject matter."[8] To deal with such cases, Quine offered what he called a "rough definition of analyticity." According to this rough definition, a sentence is analytic for a native speaker if he learned its truth by "learning the use of one or more of its words." He improved on this rough definition by "providing for deductive closure, so that truths deducible from analytic ones by analytic steps would count as analytic in turn."[9]

Quine claimed that the augmented definition accommodates such sentences as "No bachelor is married" and also the basic laws of logic. "Anyone who goes counter to modus ponens," he said, or anyone "who affirms a conjunction and denies one of its components, is simply flouting what he learned in learning to use 'if' and 'and.'" (He limits this to native speakers, he said, because a foreigner could have learned our words indirectly by translation.) Given the deductive closure qualification, he concluded that all logical truths in his sense—"that is, the logic of truth functions, quantification, and identity—would then perhaps qualify as analytic, in view of Gödel's completeness proof."[10]

In "Two Dogmas…" Quine had insisted that no statement is in principle immune to revision: revision even of the law of excluded middle had been proposed, he noted, as a means of simplifying quantum mechanics.[11] In the retrospective paper, he returns to this claim, asking "If the logical truths are analytic—hence true by meanings of words—then what are we to say of revisions, such as the imagined case of the law of excluded middle?" Echoing a question that his claim about the case of-

---

[5] Ibid, pp. 32-42.

[6] According to the conception I favor, E confirms H when E raises H's probability. Since an analytic truth has a maximal probability already, it could not be confirmed in the way Quine suggested. See chapter six, p.266. Devitt (2005), opposing the very idea of a priori knowledge on the "holist" ground that even purely logical statements must be confirmed together with other statements and "even whole theories" (p. 106) on the basis of experience, gives no hint of how the probability of "p ∨ ~p' might be raised by this process. Could it have a lower initial probability to begin with?

[7] Quine (1991).

[8] *Ibid*, p. 270. The other words quoted in this paragraph appear on the same page.

[9] The notion of closure is a mathematical one. As for analytic truth, saying that the set of analytic truths is closed under deduction is equivalent to saying that if **T** is deducible from members of this set, **T** belongs to the set as well.

[10] *Ibid.* By means of this proof Gödel showed that all the truths of first-order logic are derivable from a standard set of first-order axioms and rules.

[11] Quine (1953), p. 43.

ten prompted in the past, he raises the additional question, "Do we thereby change our [logical] theory or just change the subject, change the meaning of our words?" He answers both questions by saying, "My answer is that in elementary logic a change of theory *is* a change of meaning. Repudiation of the law of excluded middle would be a change in meaning, and no less a change of theory for that."[12]

Although Quine proceeds to say that this "more generous" view of analyticity is not really as generous as it may appear, it is important not to move on too quickly, because his rough new definition is not easy to apply. According to the new definition, a sentence is analytic for a native speaker if he learned its truth by "learning the use of one or more of its words." Of course, by the word "sentence" here Quine obviously means "sentence with a fixed interpretation." But how could one possibly learn the truth of any sentence by learning the use of one or more of its words? Exactly how could this feat be accomplished? If we do not understand this, we will not really understand the import of his rough new definition.

Since Quine said the definition "obviously works" for "No bachelor is married," this example is a good one to start with. How could one learn the truth of this sentence by learning the use of some word in it? Here is one possibility. Suppose Tommy already understands the words "no," "is", and "married." And suppose he is familiar with the grammatical structure exemplified by the sentence in question. What he does not understand in the sentence is the word "bachelor." He therefore asks his mother, "What is a bachelor, Mom?" His mother answers, "A bachelor is a man who is unmarried." How can this answer teach him that "No bachelor is married" is true? This way, I should think. The mother's utterance tells him what the unknown word applies to: it applies to any man who is unmarried. Could a man who is unmarried be married? Obviously not: No man who is unmarried is married. Since "bachelor," according to his mother, applies to a man who is unmarried, Tommy knows that no bachelor is married. He puts two and two together.

Tommy learns the truth of "No bachelor is married" in a way that recalls Kant's definition of an analytic judgment. When Kant presented his definition, he observed in passing that it could easily be extended to negative judgments.[13] The idea would be that a universally negative judgment—one of the form "No S is P"—is analytic just when the predicate concept is excluded by what is contained in the subject concept. In what way excluded? The answer is "logically excluded":[14] the ideas involved in the subject concept are logically incompatible with the predicate concept just as the ideas included in the concept of a bachelor—the ideas of being a man and being unmarried—are logically incompatible with the idea of being married. One can know that a universally negative analytic judgment is true because, on ascertaining what is contained in the concept of the subject, one will be logically assured that nothing falling under the subject concept could possibly fall under the predicate concept: the application conditions of the two concepts are logically incompatible.

I am not certain that my description of the way Tommy learns the truth of "No bachelor is married" conforms to what Quine had in mind when he spoke of learning a sentence by learning the use of some word in it. But I cannot think of another way that such a thing could plausibly be learned. Still, the pattern of this description does not apply to the way one might learn the truth of a basic law of logic. To learn the truth of "No bachelor is married" Tommy applies logic to what his mother tells him about the meaning of a word in a sentence he otherwise under-

---

[12] Quine (1991), p. 270.

[13] Kant (1997), A7, B11.

[14] The idea of a logical relation is also implicit in Kant's original definition, for he said that in affirmative analytical judgments the connection of the predicate [to the subject] is thought through [the relation of] identity. *Ibid.*

stands; he concludes that "No bachelor is married" is true because it is equivalent to "No man who is unmarried is married," and he knows that the latter is true. Evidently we do not conclude that a basic law of logic is true because something else is logically true. We do not reason in this way. How, then, are we to understand the kind of learning Quine has in mind when he speaks of learning the truth of a basic law of logic?

I really do not know the answer to this question, but a plausible candidate quickly comes to mind. When philosophers think of logic, they think of formal logic; they do so because *logical truth* is a formal notion, as is *the validity of an inference*. Today, formal logic is expounded by means of various symbols, some representing logical operations such as negation, conjunction, or universal quantification, and others representing statements and their parts—for instance, individual variables, individual constants, and relation symbols. When we learn a truth of formal logic, we learn the truth of a symbolic formula, and when we learn the validity of an argument form, we learn the validity of a symbolic pattern or sequence. Quine may suppose that we can learn the truth of certain formulas and the validity of certain symbolic patterns by learning the use of symbols contained in them.

It is convenient to begin with a valid form of inference. I have described such forms of inference as symbolic patterns or sequences; these patterns consist of statements, or premises, and a conclusion that is validly inferred from them. One of the simplest of logically valid argument forms involves conjunctions: all arguments conforming to this pattern are logically valid:

$$(p \wedge q) /\therefore p$$

To learn that this argument form is valid, we must first learn that a valid argument form is one whose proper instances have true conclusions whenever they have true premises: a valid argument form is truth-preserving. When this information is in hand, we then learn that the symbol "$\wedge$" is used to assert the truth of two statements, the two it conjoins. In learning this we learn that if a premise having the form of "$p \wedge q$" is true, both of its conjuncts are true, its first conjunct as well as its second. To learn this is to know that the form represented above is valid.

The other valid argument form that Quine mentioned is a form of modus ponens. This argument form is usually represented by a pattern containing two premises, one containing the symbol " $\supset$ " or an equivalent such as "$\rightarrow$":

$$(p \supset q), p / \therefore q.$$

To learn the validity of this form of inference we need to learn the meaning of the horseshoe symbol, " $\supset$ ". This symbol corresponds to the English "if…, then…," but its meaning is special. Its peculiarity is that it forms a conditional statement that is true whenever its antecedent is false or its consequent is true. If both premises in an argument having the form of modus ponens are true, the antecedent of the conditional premise must be true, because it is the same as the second premise. Since a horseshoe conditional is true whenever its antecedent is false or its consequent is true, the consequent of the second premise must then be true, because its antecedent is not false. But the conclusion of the argument is the same as the consequent of the second premise. Since this consequent is true, the conclusion is true. The argument form is therefore valid: when the premises are true, the conclusion is true as well. This is guaranteed by the meaning of the horseshoe symbol and the concept of a valid argument form.

A little later in "Two Dogmas in Retrospect," after expressing his generous attitude toward analyticity, Quine becomes more negative, saying "In fact my reserva-

tions over analyticity are the same as ever, and they concern the tracing of any demarcation, even a vague and approximate one, across the domain of sentences in general." By "sentences in general" he means all sentences, not just the ones expressing logical laws and truths such as "No bachelors are married." He supports this generally negative attitude with two reasons. The first is that "we don't in *general* know how we learned a word, nor what truths were learned in the process." The second is that we have no reason to expect uniformity in this matter of learning from speaker to speaker" (p. 271). Although Quine does not take these two reasons as undermining the analyticity of logical laws and examples such as the one about bachelors, we might ask why he does not. If we do not in general know how we learned a word, do we know how we learned logical words and words such as "bachelor"? And do we all learn these words in basically the same way?

The answers to these questions bring out something special about logical words (or logical symbols) and words such as "bachelor." They have, at least on particular readings, precise meanings, and they are learned in the same basic ways. Words like "bachelor" (on certain readings) are short for longer clusters of words, and when we learn their meaning—whether we are given their meaning by a teacher or parent or whether we look them up in a dictionary—we learn what groups of words they abbreviate. Like little Tommy, we learn to substitute them for their equivalents in statements that are logically true, and we thereby come to know truths that are analytic in Quine's sense. The precision of logical words has a similar result. When we learn the meaning of a logical symbol such as the horseshoe, we learn to compute the value of conditionals containing it by means of the values of the statements it connects. There is just one truth-function associated with this symbol, and when we learn what this is, we understand that symbol; we do so whether we initially encounter it in a definition relating it to negation and disjunction or in an equivalent definition that relates it to negation and conjunction. The same is true of other logical symbols. When we know what they mean, we can "by analysis" compute the truth-value of many statements in which they occur.

Analyticity, Logic, and Everyday Language

If the only truths we can reasonably claim to be analytic are those of elementary logic and trivialities such as "Bachelors are unmarried males," then the concept of analytic truth does not have the importance that empiricists take it to have. This is Quine's position, and I think he is right in holding it. I intend to provide a more satisfactory account of analytic truth in what follows, but before attempting to do so, I must first resolve some issues left over from the last chapter. Resolving these matters will bring me closer to the analysis I want to defend.

When I criticized the rationalist claim that basic logical truths can be seen to be true by a kind of direct intuition, I emphasized the extreme generality of these truths and went so far as to find instances that appeared to falsify them. I cited examples of statements that, asserting other statements to have a certain truth-value, could apparently be proved to be both true and false themselves, and I offered other examples that, owing to vague expressions contained in them, could reasonably be said to be neither true nor false and that, together with statements like them, provided apparent counter-instances to basic logical laws such as the principle of excluded middle. I even cited examples of arguments, formulated in everyday English, that some philosophers have taken to be counter-examples to modus ponens. Since these examples could not possibly be surveyed by the direct intuitions focused on general or schematic formulas that rationalists appealed to as sources for their a pri-

ori knowledge, I concluded that the rationalist's belief in the epistemic efficacy and authority of these alleged intuitions was simply and clearly unfounded.

However successful my examples may have been in refuting the basic rationalist claim about intuitive certainty, they also raise a problem for the empiricist alternative, for they raise (or should raise) serious doubts about the certain truth of the supposed logical laws that even Quine eventually described as analytic. How could we possibly know that the schematic formulas that are supposed to hold true for *all* statements corresponding to them do not, in fact, have a single falsifying instance? Do empiricists have an infallible means of surveying all instances that is not available to the rationalist? If so, what is it?

Not all empiricists would answer these questions in the same way, but one answer is this:[15] The instances to which a schematic formula is intended to apply are prescribed rather than simply surveyed. A system of logic is commonly introduced in connection with an artificial language, a system of formulas that are constructed and interpreted in specific ways. The statements of such a language system are "well-formed formulas," and rules are introduced that describe how they are properly constructed. Such formulas are interpreted by means of semantical rules, which assign semantic values to the formulas and their functionally significant parts. Possible values for the "closed" formulas of classic systems[16] are restricted to truth and falsity, and no formula can possess both these values. The kind of semantic vagueness that make it appropriate to assign an indeterminate value to particular formulas is therefore not allowed in a classical system, and one can know in advance that any legitimate formula of the system will satisfy the schemas expressing the laws of excluded middle and non-contradiction. Similarly, by placing restrictions on the kinds of predicate that can be acceptably attached to statements of certain classes, one can disallow statements such as "The sentence in the triangle is false" and make it impossible to derive in the system the sort of contradiction that I discussed in the last chapter. Thus, by playing it safe—by excluding from a logical language the sort of statement that can cause logical trouble—we can insure that classical laws are preserved there. To make this assurance maximal, to banish any possible doubt from the simplest and most trouble-free vocabulary, we can go so far as to declare that any formula leading to trouble will count as deviant all along. The system never involved an error, we may say; it was simply set up or described incorrectly.

The arguments and assertions that we evaluate in everyday life do not, of course, belong to artificial languages, and they do not consist of technical symbols that need to be assigned semantical values by technical rules. How can we use logic to evaluate them? One strategy is to adopt translations for them in a symbolic language. Thus, we might translate "If the ladder slips, the man will fall" into "L ⊃ F," taking "L", " ⊃ ", and "F" as translations, respectively, of "The ladder slips," "if," and "The man will fall." Since the formula "L ⊃ F" is easily evaluated by means of our rule for formulas containing the symbol "⊃", our evaluation for "L ⊃ F" will apply to its translation in the vernacular, "If the ladder slips, the man will fall." The acceptability of this evaluation will obviously depend on the acceptability of translating "if" in the vernacular sentence by the symbol " ⊃."[17] If the meaning of "if" in this sentence is considered acceptably close to that of "⊃," the translation will be acceptable; if not, it will not be.

---

[15] I am following Carnap (1958) here; see his chapter B.

[16] A formula is *closed* when any variable it contains is bound by a quantifier; "∃x(x is a prime number)" is a such a formula. Formulas with free variables may be "satisfied" by an object but they are not true or false.

[17] It will also depend on the acceptability of taking a certain vernacular sentence as the translation of a statement constant that must be either true or false. I comment on this below.

Another strategy for evaluating everyday arguments and assertions is to select a part of everyday language, possibly regiment it in ways that eliminate ambiguity and vagueness, and then create a logical language that is a hybrid of vernacular forms and technical symbols. A sentence of this sort of language might be "The ladder slips ⊃ the man will fall." We might even use everyday words in place of logical symbols, using "and" with the meaning of "∧" and "if" with the meaning of " ⊃ ". In this last case it will appear that we are using the language of everyday life, but we will be using just a selected part of it (not every grammatical sentence of English will count as a proper formula) and some words will not have their usual senses. To avoid paradoxes and violations of standard logical laws, we must impose restrictions on our total logical vocabulary.[18]

When Quine, "In Two Dogmas in Retrospect," agreed that the laws of classical logic and statements like the one about bachelors can be considered analytic in the rough sense he described, he left no doubt that he was thinking of logical truths as expressed in everyday language, for that is the language in which people learn the word "bachelor" as well as "if" and "and," which are the logical words he mentioned.[19] It is also clear that Quine was not thinking of the restrictions on everyday language that must be accepted if the formulas for basic logical laws are not to be falsified. (Thus, he had nothing to say about vagueness and the so-called semantic paradoxes exemplified by the statement about the false sentence in the triangle.) Also, he ignored the fact the vernacular "if" is not always used in such a way that those who have mastered its use invariably recognize the validity of modus ponens. I noted in the last chapter that the validity of modus ponens and modus tollens are, in fact, sometimes challenged by philosophers who support their case by presenting examples formulated in the language of everyday life. Now is a good time to return to the examples I presented, for they underline the importance of tying logical problems to logical systems.

The first example was this:

> If it rained yesterday, it did not rain hard (yesterday).
> It did rain hard (yesterday).
> Therefore, it did not rain yesterday.

This argument seems to have the form of modus tollens; yet the conclusion must be false if the second premise is true. It would appear that the first premise could be true. The second premise could also be true. Yet if the truth of the second premise guarantees the falsity of the conclusion, it would appear that the argument cannot be valid.

Do we have a genuine counter-instance to modus tollens? The answer is "No, particularly not if the first premise is understood as a material conditional, one that can be represented by 'It rained yesterday ⊃ it did not rain hard yesterday'." If it could be represented this way, the falsity of the conclusion would guarantee the falsity of one of the premises. It is true, as I mentioned, that both premises could be true, but reflection shows that on this interpretation they could not be true *at the same time*: they are inconsistent. This can be seen as follows. If the second premise is true at some time, the consequent of the first premise must then be false. But if this consequent is false at that time, the antecedent of the first premise must

---

[18] I emphasize the importance of this claim for current arguments about the justification of basic logical principles in Appendix 2, where I criticize some recent contentions by Paul Boghossian and Hartry Field.
[19] See Quine (1991), p. 270. Quine's view of formal logic and its relation to vernacular discourse is expounded most fully in Quine (1981).

equally be false if that premise is true. The falsity of this antecedent is therefore inconsistent with the truth of the second premise.

The second example concerned the participants in the 1980 U. S. presidential election, which was eventually won by the Republican, Ronald Reagan. Jimmy Carter, a Democrat, was second and Anderson, a Republican running as an Independent, was third. The example was as follows:

> If a Republican wins, then if Reagan does not win, Anderson will win.
> A Republican wins (=does win).
> Therefore, if Reagan does not win, Anderson will win.

Vann McGee, who discovered the example, thought it is a counter-example to modus ponens because the first and second premises seem obviously true while the conclusion seems false. Reagan won, and since he and Anderson were the only Republicans running, if he did not win, Anderson would. The conclusion seems false because the real race was between Reagan and Carter; Anderson was far behind. At the time of the election it would therefore be false to say, "If Reagan does not win, Anderson will win."

When I originally presented the example, I expressed the opinion that it is impossible to say decisively whether it is or is not an acceptable counterexample without some clarification of the English in which it is expressed. The logical word "if" featured in it is clear in some respects, but it is not clear in others, for arguments containing it can be expressed in nonequivalent symbols. Suppose we read the argument as having the following logical form:

> A Republican wins $\supset$ [$\sim$(Reagan wins) $\supset$ Anderson wins].
> A Republican wins.
> Therefore, $\sim$(Reagan wins) $\supset$ Anderson wins.

Read this way, the argument is clearly not a counter-instance, for the conclusion is plainly true: it is logically equivalent to "Reagan wins $\vee$ Anderson wins," which is guaranteed to be true if it has a true disjunct--and it does so in this case.

There are, of course, other ways of construing the argument. When I presented it as an ostensible counterexample, I suggested that the conclusion is false because the real race was between Reagan and Carter, Anderson being so far behind as to be effectively out of it. If the conclusion is read with this firmly in mind, it will appear to have a subjunctive force not captured by the horseshoe symbol. Suppose, therefore, that we interpret the "if"s in the argument as representing the counterfactual conditionality expressed by David Lewis's symbol " $\boxright$". [20] Conditionals of this kind are evaluated by reference to possible worlds or "ways the world might be." A conditional of the form "P $\boxright$ Q" is considered true just when, of all possible worlds in which P is true, Q holds in the one or the ones most similar to the actual world.

On this interpretation the argument takes the following form:

1*. A Republican wins $\boxright$ [$\sim$(Reagan wins) $\boxright$ Anderson wins]
2*. A Republican wins.
3*. Therefore, $\sim$(Reagan wins) $\boxright$ Anderson wins.

[20] See Lewis (1973), passim.

Understood this way, the conclusion is no doubt false, for in a world in which Reagan does not win but that is otherwise minimally different from the actual world (the "closest world" in which Reagan does not win), Carter presumably wins instead of Anderson. Yet the first premise is now false, and it must be true if the argument is to provide a counterexample. The closest world in which a Republican wins in 1980 is the actual world, and in this world it is not true that if Reagan were not to win that election, Anderson would. Thus, when the vernacular "if" is replaced by the technical symbol "❑→", the resulting argument also fails to provide an acceptable counter-instance to modus ponens.

Not all occurrences of "if" need be replaced by the same technical symbol, of course. Two further arguments could be obtained if one of the following formulas were put in place of 1:

4*. A Republican wins ❑→ [~(Reagan wins) ⊃ Anderson wins]
5*. A Republican wins ⊃ [~(Reagan wins) ❑→ Anderson wins]

If 1* were replaced by 4*, the result would not be an instance of modus ponens, however; for the consequent of 4* differs from 3*. If 1* were replaced by 5*, we would have an instance of modus ponens, but the first premise would not then be true. 5* is logically equivalent to the disjunction of "~(A republican wins)" and 3*, both of which are false. Thus, on these further readings we still do not have an acceptable counterexample.

Other, nonstandard readings of the vernacular "if" are possible, and it is on one such reading that Christopher Gauker defends a counterexample to modus tollens.[21] The multiplicity of possible readings of the vernacular argument raises an important question: "Just what is modus ponens?" If we do not have a particular system of logic in mind, we cannot answer this precisely. We can say that modus ponens is an argument form in which a conclusion q is inferred from a premise p and a conditional premise having p as antecedent and q as consequent; but because formulas of significantly different logical powers can be described as conditionals, argument forms of significantly different kinds can count as instances of modus ponens, some lacking counter-instances and some, for all I know, having them. The vernacular "if" is not so precise in meaning that only a single interpretation is possible for it even in a given context. If we want to single out a definite class of argument forms in speaking about modus ponens, we shall have to restrict our reference to the argument forms that can be constructed from the vocabulary of some formal system or group of systems. As I noted earlier, a "regimented" part of English may count as such a system, the precision (or logical determinacy) of its formulas depending on the way it is regimented.

It should be clear to the reader that the arguments I could confidently declare to be, or not be, counterexamples to modus ponens contain logical symbols with precise interpretations. The horseshoe symbol is not a common term whose meaning is determined by the linguistic behavior of ordinary speakers; it is a technical symbol whose logical properties are fixed by logical convention. This and other conventions permit an exact assessment of formulas whose implications are sufficiently parallel to those of certain vernacular statements to be considered the latter's symbolic transcriptions, but the vernacular statements are far less determinate in what they assert.[22] For an additional example, consider "Either something is red or everything red is green." A natural assessment of this statement is that it is a contingent truth,

[21] See above, p. 36.

[22] Frege emphasized the difference between the material conditional, "P ⊃ Q," and the "if…then" of everyday language in Frege (1962); see pp. 550ff of the reprint in Klemke (1968).

supported by the fact that red things obviously exist. But if it is interpreted as adequately symbolized by the formula "$\exists x Rx \vee \forall x(Rx \supset Gx)$," it is easily seen to be a tautology, because "$\forall x(Rx \supset Gx)$" is true if no x is R.

Analyticity Extended

If we return to Quine's rough definition of analyticity, we see that it is acceptable only on certain idealizing assumptions—that the language is appropriately "regimented," as Quine put it in *Word and Object*, that certain sentences containing "is true" and "is false" are ignored, that vagueness is disregarded or evaluated by special conventions,[23] and that logical words have the sense of certain technical counterparts. Even allowing these assumptions as trouble-free, Quine's rough definition is, as he emphasized, significantly limited and ostensibly not sufficient to accommodate the problem statements that rationalists regard as expressing synthetic a priori truths, the statements I claimed to be analytic in the last chapter.

To obtain a more encompassing definition of analyticity, it will be instructive to consider another of the definitions of analyticity that Quine criticized in "Two Dogmas…," the one focused on semantical rules. Quine actually criticized several definitions of this kind, claiming that the fundamental defect common to them all is the appeal to semantical rules: the idea of such rules is as much in need of clarification as analyticity itself.[24] Rudolf Carnap, Quine's close friend but his opponent regarding analyticity, had claimed that "the concept of analyticity has an exact definition only in the case of a language system, namely a system of semantical rules, not in the case of ordinary language…."[25] In "Two Dogmas…" Quine denied this, saying in effect that this claim puts the cart before the horse: "Semantical rules determining the analytic statements of an artificial language are of interest only in so far as we already understand the notion of analyticity; they are of no help in gaining this understanding" (p. 36).

This last remark by Quine is seriously exaggerated. As Carnap said in his *Introduction to Symbolic Logic and Its Applications*, semantical rules are rules of interpretation for what would otherwise be an uninterpreted language or formal calculus (p. 80). There is nothing obscure about the purpose of some of these rules. As I noted earlier, the horseshoe, one of the basic symbols of elementary logic, has a technical meaning that cannot be adequately explained simply by relating it to the vernacular "if" (or some counterpart in another language). To explain it adequately for the purpose of a logical system, one must specify rules of interpretation that allow us to calculate the truth-value of compound formulas containing it and other formulas. In this case the rules can be reduced to this one: A formula of the form "$p \supset q$" is true just when the formula corresponding to "p" is false or the formula corresponding to "q" is true. This is a simple, well-known rule, and to the extent that one understands it and the point of having it, one understands something about the meaning of the words "semantical rules" and the purpose of the rules they denote.

In criticizing the "semantical rule" definition of analyticity, Quine compared the notion of a semantical rule with that of a postulate. Just as no true statement is inherently a postulate, so no string of words is inherently a rule, semantical or oth-

---

[23] When vernacular discourse is regimented for logical purposes, the vagueness of everyday assertions is commonly ignored. When this sort of vagueness is explicitly recognized, a number of different logical strategies are available. One possibility is to assume a qualification that makes a vague statement sufficiently determinate to deserve a value of T or F; for example, "Tom is thin" may be read as meaning "Tom is on the thin side." For other strategies, see van Fraassen (1966), Lewis (1983), pp. 244-46, and Williamson (1994).

[24] Quine (1953), p. 36.

[25] See Carnap (1990), a short paper written in 1952 and never published by Carnap himself.

erwise. But Carnap agreed with this. In his view a particular semantical rule represents an interpretive decision, a decision about how some symbol or aggregate of symbols is to be understood in relation to the intended domain of discourse. Different decisions are always possible, but if particular decisions are made in a given case, words of a familiar kind can be used to express those decisions in that case. The same words could be used to express different decisions in a different case. As Quine said, no sentence is inherently a postulate.

The semantical value of a statement in relation to a domain of discourse is usually truth or falsity; the value of a proper name is usually a particular member of that domain; the value of a two-place predicate is a set of ordered couples in that domain; and so on. But we can also interpret some symbols by relating them to others whose interpretation is already known. Some definitions have this purpose. If "adult male" and "unmarried" are understood as belonging to the vocabulary of a regimented language-system, the word "bachelor" can be given a precise interpretation in relation to this vocabulary by the formula, "$\forall x$(x is a bachelor $\equiv$ (x is an adult, male human being $\wedge$ x is unmarried))."[26] The idea would be that regardless of the meaning that the word "bachelor" might have in everyday language, in the context of the regimented system it is to be understood as an abbreviation of the words appearing in the right-hand side of the defining formula. One may wish to introduce a strict sense of "bachelor" if a strict sense is needed for special purposes.

Carnap was convinced that a precisely specified language system is needed for the concept of analyticity because he thought words have no "clearly defined meaning" in ordinary language. It is easy to miss the reasonableness of his view here. Consider "bachelor," a word for which a strict sense might conceivably be needed. In everyday life the word is not only ambiguous, but it is often used quite loosely. As for ambiguity, the word is now occasionally applied to young women living alone or to people possessing a B.A. or B.S. degree (see the OED); as for looseness, people are actually apt to disagree (as Gilbert Harman observed) about whether the word is applicable to the pope, who is not married in any ordinary sense, or whether it should be applied to a man who has lived with a woman for several years without getting married.[27]

In view of the controversy about truths that epistemological rationalists claim to be synthetic a priori, it is worth considering an example that arose in a dispute between Carnap and Quine on analyticity. The example was "Everything green is extended," which Quine said he hesitated to classify as analytic because of an incomplete understanding not of "green" or "extended" but of "analyticity." Carnap said it seemed "completely clear" to him that the difficulty lies in the unclarity of "green," which betrays an indecision whether to apply the word to a single space-time point.[28] "Since one scarcely ever speaks of space-time points in everyday life," he said, "this unclarity about the meaning (or intended application) of 'green' plays as small a role [in everyday life] as the unclarity about whether the term 'mouse' should also be applied to animals which, apart from their greenness, are completely similar to the mice we know, but are as large as cats."[29] This lack of clarity is unimportant for the practical purposes of everyday life, but it is vitally important for the philosophical question about the analyticity of "Everything green is extended." To

[26] Technical definitions have the form of a biconditional or an identity statement. See Suppes (1957), Ch. 8, "Theory of Definition."
[27] Harman (1996), p. 399.
[28] In his language form IIB described in Carnap (1958), Carnap defined space-time points as "the smallest non-empty spatial regions; see p. 160.
[29] Carnap, "Quine on Analyticity," p. 427.

settle the latter, Carnap thought, we must make our meaning of "green" or color words generally more precise in relation to our thought about points.

The idea of making one's meaning more precise in certain respects, or in some respects rather than others, was very important for Carnap and is, I believe, very important for the subject of analyticity. Carnap first called attention to the importance of a partial analysis in 1936, when he wished to introduce predicates for dispositions into the context of a technical language having the horseshoe as its sole symbol for conditionality. He could not define "x is water-soluble" by the conditional "x is immersed in water ⊃ x dissolves," because, owing to the truth of material conditionals with false antecedents, anything never immersed in water would then count as water-soluble. To avoid this difficulty, he introduced the idea of a "bilateral reduction sentence," a formula by which the meaning of a disposition predicate is specified incompletely, only for instances in which the relevant test condition is satisfied. The general form of such a reduction sentence is "$Q_1 \supset (Q_3 \equiv Q_2)$," where "$Q_1$" and "$Q_2$" represent preexisting predicates of the scientific language and "Q_3" represents the predicate whose meaning is being specified for cases in which the test condition "Q_1" is satisfied.[30] Applied to the predicate "water soluble," the reduction sentence lays down a necessary and sufficient condition for the application of this predicate to objects immersed in water. The predicate's application to objects not so immersed would remain undetermined in basically the way that the application of "is bald" is undetermined for cases in which a person showing a lot of scalp still has a significant amount of hair.

The practice of reconstructing the meaning of vernacular words, which I discussed in chapter one in connection with David Lewis's treatment of "S knows that P," Carnap called "explication." When the meaning of a word or formula is fully explicated, or completely reconstructed, it is introduced into technical language by explicit definitions whose *definiens* consist of words or symbols whose meaning is antecedently clear and unproblematic. For cases in which the meaning is explicated only incompletely, Carnap first used the label "meaning postulate" and later changed it to "A-postulate":[31] for him, A-postulates are the formulas providing the partial explications. These explications are not generally intended to specify some part or aspect of the meaning that a word or group of words already possesses; they are used to stipulate the meaning they have in a specified (or tacitly understood) context: either the context of a technical language or discourse, or that of some discussion.

Carnap illustrated the point of an incomplete stipulation in a paper called "Meaning Postulates."[32] Suppose a person constructing a certain system wishes to use the symbolic predicates "Bl" and "R" in a way corresponding to (but not necessarily the same as) the way "black" and "raven" are used in everyday life. Speaking of such a person, Carnap says:

> While the meaning of 'black' is fairly clear, that of 'raven' is rather vague in the everyday language. There is no point for him to make an elaborate study, based either on introspection or on statistical investigation of common usage, in order to find out whether 'raven' always or mostly entails 'black.' It is rather his task to make up his mind whether he wishes the predicates 'R' and "Bl" of his system to be used in such a way that the first logically entails the second. If so, he has to add the postulate (P_2) '(x)(Rx ⊃ Bl x)' to the system, otherwise not" (p. 225).

---

[30] See Carnap (1936 and 1937).
[31] See Carnap (1966), p. 261.
[32] Carnap (1956) pp. 222-229.

If the postulate P2 is added to the system, the person constructing it has thereby stipulated how, in the context of the system, the predicate "R" is to be understood in relation to a symbolic predicate corresponding to "black." If "R" is applicable to a thing x, "Bl" must be applicable to it as well.

Gilbert Harman once said, "…stipulative definitions are assumptions. To give a definition is to say 'Let's assume for the time being that the following equivalence holds'."[33] This is wrong. Assumptions can be false; stipulative definitions cannot.[34] If I decide to use "raven" in accordance with the stipulation (holding for a certain context) that nothing non-black will count as a raven, I will not be proved wrong if something that might be called a raven in the ordinary sense--a bird indiscernible from a raven except for being white--should be observed. It would simply not be a raven in my stipulated sense. Using my special terminology, I might call it a "waven" and say that ravens and wavens in my sense of the words are pretty clearly subspecies of a distinct kind that might be called "dravens." Seeing such a bird might move me to bring my special terminology more into line with common usage and to use "raven" as people ordinarily do. But I would not have made an error in using "raven" as I formerly did.

A meaning postulate, as Carnap understood it, is very close to the sentences featured in the "modest" sort of analytical account that Williamson offered for the concept of knowing.[35] This kind of account discloses the conceptual connections between a target concept and certain others, and in doing so it provides a kind of non-reductive analysis of the target concept. In explaining how knowing can be understood as being the most general "factive, stative attitude," Williamson identified a number of analytic implications in which "knows" participates. Three obvious examples are the following:

If S knows that P, it is true that P.
If S remembers that P, S knows that P.
If S sees that P, S knows that P.

Carnap differs from Williamson in having serious reservations about the precision and determinacy of everyday language. As I have explained, his postulates are to be understood as stipulations rather than complete or partial analyses of existing usage. He generally expected them to reflect existing usage if there is no need, scientifically or philosophically, to diverge sharply from it; but he thought that we are bound to diverge in some degree if we wish to be clear and precise.

Although I am somewhere between Carnap and Williamson in my attitude toward everyday language, I have no doubt that Carnap's strategy of providing stipulative explications allows us to introduce a broader sense of analytic truth than the one given by Quine's "rough definition." Statements so explicated are analytic *for us* (not analytic generally) because they represent part or (conceivably) all of what we mean in using the words they contain. People who speak "our language"—people who speak English, for example—need not mean what we mean by every word, and our explications need not be valid for what they say. This broader sense of analyticity does not therefore identify the analytic sentences of a whole natural language or dialect, though we may wish and even recommend that others adopt our usage in preference to theirs. Nevertheless, this limited and local conception of analyticity is

[33] Harman (1996), p. 399.
[34] They can, of course, be revised, abandoned, and the like. But revision and so forth is not the same as falsification. In Appendix 3 I discuss some conditions that an acceptable stipulation must satisfy.
[35] See chapter one, p. 00.

sufficient for epistemology. It allows us to dispose of the issues rationalists raise by means of the problem examples I discussed in chapter two.

Consider again the statement, "Nothing can be both yellow and green all over at the same time." As my discussion in chapter two made clear, this statement need not even be true. "Yellow" and "green" are highly generic predicates that are not used in exactly the same way by all speakers of English. Although they are perhaps normally regarded as incompatible, they can be used, as Harry of my story did, in a way that makes them jointly applicable to the same part of a leaf or shrub. Something with the determinate shade Harry called "green-yellow" may be described as both green and yellow all over, for both colors are there, all over. What are clearly incompatible are determinate color shades: If something is green-yellow in Harry's sense or yellowish-green in Mary's sense, it cannot at that time also have any other determinate shade of color. This incompatibility is not a matter of ontological fact that is independent of classificatory conventions; it is a consequence of how we individuate a thing's specific color at a time. We *could* restrict ourselves to a purely generic means of attributing colors, calling things either yellow, green, red, or blue, and so on; and if we did so, there would be no definite error in our describing something with Harry's green-yellow shade (which we would not then distinguish as such) as both green and yellow at the same time.

In discussing color incompatibility in the last chapter, I said that we do in fact identify specific colors in a way that assumes indiscernibility as an identity condition for them. We consider a determinate color A to be the same as a determinate color B just when A and B are indistinguishable.[36] When we conceive of specific colors this way, we are tacitly accepting a convention that renders it analytic for us that nothing can have two different determinate colors at the same time.[37] The analyticity here is not peculiar to just a few of us; it holds for all who accept the convention—all who identify specific colors this way. Many of the tacit conventions that render statements analytic for members of a group govern aspects of the use of words or sentences that are as wholes vague or hard to define. It is not easy to say exactly what a fake object is, but there is no doubt that if something is a fake duck, it is not a real one, and there is no doubt that that if Nero fiddled while Rome burned, Rome was burning while Nero fiddled. Grammatical structures that do not appear in formal languages also warrant inferences that are valid for those who use them. If someone says of a friend, "Lacking an umbrella, she hit him with a shoe," we are normally entitled to infer that, if the speaker is right, the hitter lacked an umbrella, hit a person or animal with a shoe, and did the latter because of the former. The truths of these conditionals and the acceptability of this last inference are not ascertained in Quine or Frege's way, by making deductions from logical truths and accepted definitions; they immediately come to mind as the consequence of tacit conventions accepted by all who use the relevant language in a normal way and can think abstractly about truth and validity.

At the present time 350 million people speak English as their first language and around 450 million speak it as a second language.[38] These people live in different parts of the globe, have conflicting interests and customs, and vary greatly in education and general knowledge. Generalizing about the structure of English or the meaning of this or that English word is therefore inherently risky. The same is true, of course, for any other widely used language. Realizing this, I am tempted to say

---

[36] A more satisfactory of expressing this is to say that x and y (or regions on their surfaces) have the same determinate color just when *they* are indistinguishable in color. The point of this observation will become evident in chapter 4.

[37] See the proof given in footnote 46 of chapter two.

[38] Ferguson (2002), p. 304.

that we can justifiably speak of analytic truths only when we can relate them to logical systems and explicit stipulations, the latter being either complete or partial. But this attitude is really too cautious. The examples I gave in the last paragraph make it obvious that words, phrases, clauses and constructions in existing dialects of natural languages have implications so vital to the meaning of what they are used to say that any alert and attentive speakers of a relevant dialect would find it odd, puzzling, or paradoxical to question them. When this condition is satisfied by a word or symbol, it seems to me that a sentence of the dialect clearly and unambiguously expressing an appropriate implication can reasonably be regarded as analytically true for those alert and attentive speakers.[39]

In making this last claim I am obviously adding to the conception of analyticity that Carnap offered. I am not limiting analytic truths to statements that are true for certain speakers by virtue of explicitly identified semantical rules and complete or partial stipulative explications; I am also including statements whose truth is ensured by the conventions that those speakers tacitly apply in making them—conventions whose implications are so vital to the meaning of the words and structures being used that the speakers would find it odd, puzzling or paradoxical to question them. These latter statements can, of course, be related to the sort of semantical rules and complete or partial explications that Carnap described. The procedure is this: If explicit semantical rules and complete or partial explications sufficient to demonstrate the truth of those statements *were* formulated, brought to the attention of the relevant speakers, and satisfactory explained to them, the speakers *would* then accept them as making explicit the meaning they attach, wholly or partly, to the words, phrases, and constructions involved in those statements. If the speakers would not do this, and if no alternative explanation of their negative attitude were available, the statements in question could not reasonably be regarded as analytic for them there and then.

The meaning speakers attach to the words they use in saying this or that need not be associated with a dialect in a narrow sense of the word. This is an important matter, because the speakers might comprise a very small group, even a singleton, adopting special conventions for a particular publication or a serious conversation. Just the other day, in a discussion with another philosopher, I temporarily adopted a special convention for the word "variable." Because adjustments and qualifications pertinent to a person's usage are often partial, temporary, and relevant to just this or that audience, a satisfactory account of analyticity should always be related to some reasonably determinate context. The explicatum should be "Φ is analytic for Σ in context X," where Σ is a class that includes the relevant persons (the speakers and hearers, or just the speaker or speakers) and X includes the parameters identifying the context. S may be analytic for Tom and Sally in the context of a particular discussion; S' may be analytic for me in the context of a book or chapter I have written; and so on. The relevant explicans (or analytical account) that provides the explication should *ideally* list the relevant semantical rules and the full or partial explications that characterize the conceptually determinate aspects of the language used by the persons Σ in the context X. In practice this is an excessively demanding requirement for speakers of natural dialects, because they are normally accustomed to relying on tacit conventions that only experts can be expected to identify and de-

[39] The analytic character of the informal inferences normally involved in evaluating formulas and argument forms by reference to semantical rules is to be understood along these lines. The inferences could, of course, be formalized, in which case their validity could be assessed by higher-order rules. But the assessment would not *make* a formula tautologous or an inference valid. "P $\vee$ ~P" is a tautology if it is an instance of a schematic formula all of whose proper instances are true. There are different ways of discovering whether "P $\vee$ ~P" has this property; one is by using a truth table.

scribe.[40] But special meanings should nevertheless be clarified in this way. If the meanings are special, they are usually not associated with tacit conventions.

Some Examples and Arguments by Kripke

Having explained Carnap's approach to analyticity and my extension of it, I can now attempt to come to terms with some important unfinished business-- specifically, the examples illustrating the alleged necessity of a thing's origins that Kripke mentioned in two footnotes of his *Naming and Necessity.* I discussed these examples in the last chapter. I noted that Kripke said one of the examples is "susceptible of something like a proof," and reflection convinces me that the argument he seemed to have in mind for this example can be converted into arguments that apply to the others. These arguments depend on an axiom of modal logic that, like any logical axiom, is arguably analytic in Carnap's sense of being true by virtue of semantical rules. If the arguments succeed, the examples can then be considered analytic in the sense I have explained; they will provide no support for epistemological rationalism.

The arguments I shall consider make use of a strategy Kripke included in a footnote to the second edition of *Naming and Necessity* to support the principle that if an object has its origin in a certain hunk of matter, it could not have had its origin in any other matter. After formulating this argument in way that makes its logic easy to follow, I will show how it can be revised to support the other examples.

The principle to be proved by the first argument can be stated as follows. If M1 had its origin in a hunk of matter H1, then M1 could not have originated from any hunk H2, where H1≠H2. This principle is intended to hold for all M1, H1, and H2; the argument, in showing that it holds for any arbitrarily chosen values of these variables, shows that it holds for them all. The argument proceeds by conditional proof. Assume that there is a possible world in which M1 had its origin in H1 (as in the actual world) and that an object very like M1 was made from a different hunk of matter H2. Since H1 and H2 are distinct hunks of matter, M1 is distinct from M2 in this world. But if two objects are distinct in any possible world, they are distinct in every possible world. This is a theorem of Kripke's modal system. Yet if M1 could not be identical with M2, which represents *any* relevantly similar object made from a different hunk of matter H2, M1 could not have originated from any such hunk. Since an origin would be impossible.

I said above that the argument just given could be adapted to provide arguments supporting the other principles about the necessities of a thing's origins that Kripke discussed. Take the principle about parents: If C's biological parents are P1 and P2, C could not have been born to anyone other than P1 and P2. To prove this, assume that there is a possible world in which C's biological parents are P1 and P2, as in the actual world, but that a person D, just like oneself otherwise, was born from other parents, P3 and P4, at the very same time. In this world, clearly, D ≠ C, since they have different biological parents. By the kind of modal reasoning given in the last paragraph, it follows that there is no possible world in which D = C. Since D is representative of *any* possible person born of different parents when C was born from P1 and P2, it follows that C could not have been born from different parents.

Kripke says (p. 114n) that the two arguments I have just given ultimately rest on a modal principle that he calls "the necessity of distinctness."[41] The principle described by these words is usually expressed by saying that if **a** and **b** are distinct things—that is, if **a** ≠ **b**—then it is necessary that **a** ≠ **b,** but the argument I have given above requires a stronger principle—namely, that if it is possible that **a** and **b**

[40] Consider the conventions for irregular verbs described in Pinker (1999).
[41] Kripke (1980), p. 114.

are distinct, then it is necessary that they are distinct. Expressed differently, the principle is that if there is a possible situation in which **a** and **b** are distinct things**,** then they are not be the same thing in any possible situation. It seems to me that this principle accords with what we mean in speaking of the same and different things. If there is a possible situation in which I am distinct from some other person, how could I possibly be myself and also be that person in some other situation? The meaning of vernacular words is not decisive for logic, of course, but the operators of Kripke's modal system are sufficiently parallel in meaning to their vernacular counterparts to be used in their place in a complicated argument. Since the semantical rules of his system allow us to prove that the strong necessity-of-distinctness principle is logically true, we can justifiably regard it as analytic in Carnap's sense. And since the arguments I have reconstructed depend on that and other logical principles, the claims Kripke supported by means of those arguments—if the arguments are in fact satisfactory—deserve to be regarded as analytic as well. No patently synthetic a priori premises are needed in their defense.

I added the qualification, "if the arguments I have given are in fact satisfactory," because I do not believe that they actually prove what they are intended to prove. Take the second argument, which is intended to prove that if the biological parents of a person C are P1 and P2, C could not have been born to anyone other than P1 and P2. To prove this, the argument supports the principle that if it is possible for a person actually born of parents P1 and P2 to have those parents when a very similar person has other parents, then the first person could not be identical to (or one and the same as) the second person. But this last principle is not equivalent to the principle the argument purported to prove, nor does it entail that principle. This is evident from the fact that the argument relies on the possibility of two very similar persons, C and D, with different parents coexisting in a possible situation. The fact that C is not identical with D in this situation does not show that there is no other situation in which C has the parents D has in this situation. If C had those parents in some other situation, C would not have his (or her) actual parents there, but C would still be himself (or herself), not some other person.

Although I am not convinced by the arguments I have considered, I would not insist that the principles about the necessity of origins that Kripke discusses are in fact false. I offer no opinion on that subject. I will say, though, that if those principles can be proved by some argument,[42] the argument will be analytical and the principles will be shown to be analytic. There is no plausibility in the idea that they are intuitively obvious or deducible from premises that are not analytic in the extended sense I have introduced in this section.

Beliefs, Propositions, and Analyticity

What I have been saying about analyticity in the last two sections supports a language-centered account of the subject. Can it be extended to accommodate the apparent fact that judgments and beliefs may also be analytically true? These psychological states may be expressed linguistically, or put into words, but they are evidently not themselves linguistic entities and it would appear that they are as susceptible of a priori truth or falsity as any statement. Stipulations about the meaning of words can hardly be pertinent to their falsity or truth. Or so it would seem.

To evaluate this important objection, it is vital to have a defensible conception of a judgment or belief, to know just what they are and how they are put together. Someone new to epistemology might think that the nature of these states is obvious to any thinking person, but the reality is quite otherwise, at least if we go by what

[42] Nathan Salmon discusses other arguments for Kripke's conclusion in Salmon (2005), ch. 7.

philosophers say about them. A very common claim is that judgments, beliefs, doubts, suppositions and a host of other propositional attitudes consist in some relation to a "proposition."[43] A judgment is always a judgment that P (for some P); a belief is always a belief that P; and analogous claims hold true for the other attitudes. What is common to them is some proposition or other; they differ in the way they are related to a proposition. Believing and doubting involve relations that are virtual opposites; believing and suspecting are similar in some respects but different in others; believing and opining are substantially the same.

If we are to take this view of so-called propositional attitudes seriously, we have to know what a proposition is. The classic view of such a thing, the one worked out in what David Kaplan called "the Golden Age of Pure Semantics," was introduced by Gottlob Frege and refined by Rudolf Carnap.[44] Frege viewed a proposition as the "sense" or meaning of a sentence. Since the words of a meaningful sentence are themselves meaningful units that contribute to the meaning of the whole, the sense of a sentence is a function of the meanings (or senses) of its words. According to Frege, the names and predicates of a sentence have "concepts" as their senses, [45] and these concepts may be singular as well as general. Consider the sentence "Socrates is wise." Corresponding to the descriptive words in this sentence are two concepts, the individual concept corresponding to "Socrates" and the general concept corresponding to "wise." Frege's account of the relation between these concepts is somewhat confusing; in one place he appeared to describe it as a relation of subordination (of the individual concept to the general one).[46] Carnap described it as attribution or predication: presumably the general concept is predicated of the individual to which the individual concept applies.[47]

How are propositions so understood related to believing? Frege and Carnap appear to differ on this matter. Judged by his essay, "The Thought: A Logical Inquiry," Frege seemed to believe that propositions can be directly apprehended and so accepted by the believer independently of any sentence. For him, the basic relation between person and proposition was one of "apprehension."[48] Carnap, by contrast, held that our access to propositions involves the use of sentences. In his view, the statement "John believes that P" has the sense, approximately, of "John is disposed to an affirmative response to some sentence that expresses the proposition that P." [49] Since Carnap himself applied the predicate "is L-true" (his equivalent for "is analytic") only to sentences for which semantical rules have been given, anyone who accepts his analysis can apply "is analytic" to propositions only in some extended sense. The only plausible way of doing this is to say that a proposition is analytic if it is expressed by (or is the intension of) a sentence whose truth, in a system S, can be calculated on the basis of the semantical rules of S alone. The treatment of analyticity I outlined in the last sections obviously accommodates this strategy very well. One's belief is analytic just when the sentence to which one is belief-related is analytic in the sense I specified.

---

[43] I comment on Scott Soames's version of this view in 9.

[44] Kaplan (1991a), p. 214.

[45] This common interpretation does not accord with some of Frege's explicit claims. In "On Concept and Object" (Frege [1892]) he described a concept as the "nominatum" of a predicate and perhaps considered the corresponding sense as the "mode of presentation" of this concept. Carnap (1956) says that the interpretation I assume here, which he and Alonzo Church accept, "is in accordance with Frege's intentions when [as he occasionally does] he regards a class as the (ordinary) nominatum of a…common noun and a property as its (ordinary) sense" (p. 125).

[46] See Frege (1892), p. 48.

[47] Carnap (1956), p.

[48] Frege (1965), p.307.

[49] Ibid., pp. 54-62.

Frege's view of our access to a proposition is obviously far more attractive to a rationalist than Carnap's, but the relative merits of these views are no longer very significant since the classic view of propositions involved in them has been seriously undermined by recent work in semantics. The fundamental defects of the classical view can be traced to proper names whose supposed correlates in a proposition were taken to be individual concepts. Frege and Carnap thought that these concepts were needed to connect names to objects in the world, but the required individual concepts do not, in general, exist: there is no generally shared conceptions that single out the referents of commonly used names, and historical individuals such as Socrates and Aristotle may fail to satisfy the descriptions that people commonly associate with them. The connection between proper names and their referents is now generally thought to be "direct" rather than mediated by some associated concept. A connection is set up in a community by various talk and behavior, sometimes by acts of naming or dubbing, and the name is then spread through the community of language-users by talk and actions, moving from "link to link as if by a chain."[50] No individual concept, no uniquely identifying description, is needed in this process.

Demonstrative expressions such as "I," "here," "now," "he" and "she" are also not connected to their referent by some individual concept; they too directly refer to their referent. They have, it is true, as David Kaplan has emphasized, a distinctive character by means of which speakers and hearers can identify their referent in this or that context of utterance, but there are no propositional components, no concepts, that single out those referents. As a matter of fact, auditors will commonly interpret an utterance containing demonstratives by different words, even when speaker and auditor share the same language. I say "The book is here on this desk," and my hearer interprets me as saying that the book is there on that desk. Mary tells me "I will meet you on that corner tomorrow," and the next day I, waiting on the right corner, think, "She said she would meet me on this corner today." Her assertion and my thought of what she said have no common, classically conceived propositional object.[51]

What conception has emerged from the breakdown of the classical conception of propositions? No single conception appears to be dominant.[52] Some philosophers who once accepted the classic conception have simply given up on propositions altogether.[53] Others have retained classical propositions for fully general sentences but developed new conceptions for sentences containing proper names and demonstratives. One thing common to leading conceptions of singular propositions—the propositions expressed by atomic sentences containing proper names—is that the referents of the names are said to exist *within* those propositions. David Kaplan, who was the first to develop a view of this kind in the late twentieth century, cited the early Bertrand Russell as his precedent. Kaplan himself described the proposition expressed by "Socrates is wise" as the ordered pair consisting of the man Socrates and the property *Wise*—that is, as <Socrates, *Wise*>.[54] Other philosophers have described singular propositions in other ways, but they have retained the Russell-Kaplan strategy of "loading" referents into these propositions.[55]

The truly revolutionary features of propositions so understood is that they are not themselves objects that represent the world, as classical propositions were, but

---

[50] Kripke (1980), p. 91; see also pp. 92-164.

[51] See Perry (1979). Another serious problem with classically conceived propositions is presented in Kripke (1979).

[52] See King (2001) and Fitch (2002).

[53] This is Chisholm's response; see Chisholm (1997),

[54] Kaplan (1991a), p. 221.

[55] For a helpful discussion of specimen examples of these alternatives see the article "Propositions" in the *Stanford Encyclopedia of Philosophy.*

helpers or interpreters (it is hard to say which) of other objects—namely, sentences—that do represent it. Such propositions are often informally referred to as "what is said"[56] by utterances of sentences in various contexts, but this way of speaking is not really appropriate. If I say that Tom Smith has a silly smile, I say something *about* Tom Smith; the man himself is not happily described as part of what I say. In fact, if the "property" that Kaplan takes to be the second component of the singular proposition <Socrates, *Wise*> is the sort of thing that can exist in the world (as many philosophers suppose) neither component of this proposition is reasonably considered a part of what someone might say. Both parts are rather things one may refer to or talk about in saying this or that.

Philosophers concerned with the semantics of counterfactual conditionals and statements of necessity and possibility commonly advocate another fashionable conception of propositions.[57] According to this conception, propositions are either sets of possible worlds or functions from possible worlds to truth-values. But sets of possible worlds can hardly be grasped by the mind in the way Frege and others thought propositions could be grasped, and the same is true of functions from worlds to truth-values, which are commonly viewed as sets of ordered couples, each couple consisting of a possible world and an associated truth-value, specifically truth. This conception is obviously quite technical, but it is not really hard to understand, and it has the merit, from my point of view, of being entirely compatible with the view of analytic truth that I developed in the last section. I want therefore to say some more about it here.

Consider the sentence "Bachelors are unmarried." Understood in the usual strict or idealized way, this sentence is true in a wide range of possible worlds or "ways the world might be."[58] Conceived of as a function (or many-one relation)[59] from possible worlds to truth-values, the proposition expressed by the sentence "Bachelors are unmarried" is the function that assigns the value T (truth) to a world just in case the bachelors in that world are unmarried. Conceived of more simply as a set of possible worlds, the proposition is the set of worlds in which all bachelors are unmarried. But which worlds are in this set? Or, equivalently, which worlds are assigned the value T by the relevant function? The answer is "All possible worlds whatever." How do I know that this answer is true? Because the sentence "Bachelors are unmarried," understood in the usual strict way, is analytically true. Any person in any possible world that counts as a bachelor is guaranteed to be unmarried. This is owing to the meaning of the predicate "bachelor" or to any predicate that properly translates it. We do not have to examine the contents of a possible world to know that it is assigned the value T by the bachelors-are-unmarried function. We know this by knowing what a bachelor (in the sense in question) is supposed to be.

It is worth observing here that the possible-worlds conception of a proposition moves the notion of a proposition away from the classical conception for a reason I have not yet mentioned. According to the classical conception, a proposition is the fundamental bearer of truth: it is what is true in a fundamental sense. As Frege put it in a famous passage: "What does one call a sentence? A series of sounds; but only when it has a sense…. And when we really call a sentence true, we really mean its sense is."[60] (Recall that a proposition, for Frege, is the sense [*Sinn*] of a sentence.) But a function from worlds to truth-values or a set of possible worlds is not

---

[56] See Appendix 4.

[57] Se Lewis (1973), p. 46f.

[58] David Lewis describes possible worlds as "ways the world might be." See Lewis (1986), p. 2.

[59] A relation R is said to be many-one just when for every object x in its domain (the entities it relates to something) there is just one object y in its range (the entities it relates something to). *The biological father of* is thus a many-one relation, since everyone has just one such father.

[60] Frege (1965), p. 290.

really a bearer or possessor of truth; it is not itself true at all. This point, oddly enough, seems to be overlooked even by philosophers who actually make it. In his excellent encyclopedia article on propositions, Jeffrey King says this:

> Intuitively, it [the intension of a sentence, a proposition] maps a world to the value true if the sentence is true at that world. Thus the intension of a sentence can be seen as the primary bearer of truth and falsity at a world: the sentence has the truth value it has at the world in virtue of its intension mapping that world to that truth value.[61]

What King actually says to be true here is a sentence, or possibly a world; the proposition is a "bearer" of truth only in the metaphorical sense that it "carries" (maps) the world to a truth-value. Thus, propositions on this conception not only fail to be "what is said"; they are no longer even true or false.[62]

In view of the general failure of the classical conception of propositions, it is important to consider an alternative to the attitude-object view of propositional attitudes, the one that describes them as relations to a propositional object. The standard alternative, historically speaking, is known as conceptualism, the view held by such philosophers as Kant. According to this view, propositional attitudes—believing, judging, supposing, and so forth—have "contents" rather than "objects." The content of a thought that Socrates is wise has two principal constituents. The first constituent is a singular idea, one that represents Socrates in the way that the name "Socrates" represents him. Following Kaplan, we can say that the idea represents him *directly*. The other constituent is a general idea, a concept in Kant's sense, one by means of which the referent of the subject idea is characterized as wise. This is substantially Kant's account of the matter, though his logical apparatus is simpler than what we would use today. Since the predicate concept in this last case is not contained in the subject (it could not be, since the subject has the character of a name) Kant would declare it to be synthetic. If the matter were otherwise—if the predicate were so contained—it would be analytic.

How would this conceptualist account of thought relate to my extended account of analyticity, the one involving semantical rules and complete or partial explications? This way: Just as such rules and explications tell us what reality (or an item of reality) must be like if a certain word or formula is applicable to it, so analogous rules and explications tell us what reality must be like if the idea or thought expressed by a given formula is applicable to it. The process by which we fix the content of an idea or thought is fundamentally the same as the process by which we fix the meaning of a word or formula: we adopt appropriate principles of reference, equivalence, and inference. To make my idea of a minimal if-then relation clear, for instance, I can adopt the principle that a compound thought involving this relation will be true just when either its antecedent is false or its consequent is true. In adopting this principle I am clarifying my idea. I can also form ideas in a deliberate way. The fundamental fact is that thoughts, as conceptualists describe them, have the semantic properties of the words that "express" them, and their content is determined by corresponding principles.

Philosophers who believe in classically conceived propositions say that they can be expressed in different languages by sentences that are good translations of one another. Since a claim of this kind is not intended to be a tautology, it should be possible to say, at least in a general way, what counts for mutual translatability

---

[61] King (2001).

[62] It is interesting to note that Carnap, in a book where he espoused a basically Fregean view of language, insisted that "truth in the semantical sense is a property of sentences." See Carnap (1956), p. 93.

without introducing the idea of a proposition. If I were asked how we could identify two sentences that have this status on a standard reading, I would say that each sentence should be built up in equivalent ways by words that apply to the same objects and have corresponding implications. To have corresponding implications the words would have to be such that, if they are applicable to certain objects, additional words that are good translations of one another are also applicable to those objects. The syntactical and semantical similarities that must exist here are very complicated, but if we can ascertain that they do exist, we will know everything we have to know to decide if they are good translations or not. We will not have to appeal to anything abstract that they both "express."

If propositions, understood as classically conceived abstract objects, are not needed to account for the translatability of one sentence by another, and if thoughts and statements can be semantic counterparts without being related to a common object of this kind, then such propositions are not really needed for an acceptable semantic account of either thought or speech. Since I accept both antecedents of this last conditional, I accept its consequent: classically conceived propositions are not in fact needed for a semantical analysis of either thought or speech. This conclusion is supported, as Kaplan and others have emphasized, by the directly referential character of names and demonstratives, but it is also supported, as I shall argue in the following chapter, by the directly attributive character of predicates. As they are used in properly formed sentences, predicates can be used to characterize or describe objects without relating them to a further object, a property in one sense of the word, that somehow does the job for them.

Near the end of the "Afterthoughts" that he appended to his paper "Demonstratives," David Kaplan expressed the view that "our connection with a community in which names and other meaning-bearing elements are passed down to us enables us to entertain thoughts *through the language* that would not otherwise be accessible to us." We become capable of thinking about things in the world as the result of having experienced various things ourselves, but we also gain the capacity "vicariously," he said, "through the symbolic resources that come to us through language. It is the latter—vocabulary power—that gives us our apprehensive advantage over the non-linguistic animals."[63] I agree with these sentiments completely. We do gain the capacity to think about occurrences in ancient history, exotic forces and fields in subatomic physics, and even certain traits and quirks of everyday acquaintances only by means of words we learn from parents and teachers, textbooks and dictionaries, newspapers and television. We do not master perfectly the words we accumulate, and our sources are also imperfect transmitters of collective verbal wisdom, so there is usually some lack of fit between our speech and thought and the speech and thought of others. For this reason, there are not many words in common use with the precise univocal meanings that could justify the definitions and analyses that many philosophers construct—if those definitions and analyses are not partial or explicative in the sense I have explained. An analytic/synthetic distinction is not really possible for the whole of our language as it actually is. But clarifications and reconstructions are always possible. These provide the basis for an acceptable, philosophically useful account of analytic truth.

[63] See Kaplan (1992b), pp. 603f.

Chapter Four

PROPERTIES AND CONCEPTS

The argument of the last chapter made it obvious that epistemology is closely connected with metaphysics. Although I discussed several conceptions of propositions there, both classic and very recent, I had very little to say about properties, the entities rationalists such as Chisholm and BonJour claim to apprehend directly and regard as their source of synthetic a priori knowledge. I have done my best to refute the rationalist position, but I cannot realistically hope to create conviction in my alternative if I do not come to terms with the nature and reality of properties. The importance of doing this is amplified by the fact that no single conception of properties is clearly dominant in current discussion. Two conceptions appear to be primary rivals, and what is plausible on one of them is implausible or even nonsensical on the other.

What are Properties?

Properties have historically been conceived of in two fundamentally different ways. One is basically Platonic and the other Aristotelian. I say "basically" here because I am thinking of generically similar conceptions, instances of which may differ significantly from the conceptions actually held by Plato and Aristotle. Philosophers holding a basically Platonic conception sometimes speak of their properties as "concepts." One naturally thinks of concepts as general ideas, as Kant did, [1] but there is a precedent going back to Frege for thinking of concepts as objective entities that particular things may exemplify or be subsumed under. In the essay "On Concept and Object" Frege actually says "I call the concepts under which an object falls its properties,"[2] and Bertrand Russell confidently asserted that "awareness of universals is called *conceiving*, and a universal of which we are aware is called a *concept*."[3] Philosophers in the other tradition, the one I am calling "basically Aristotelian," conceive of properties as entitles that are literally present in the world around us; [4] they are either constituents of particulars or attached to them in some intimate way.

Both of these traditional views are still defended, or at least espoused,[5] but they are accompanied by two more, one of which is technical and idiosyncratic. The technical one is accepted mainly by writers on the semantics of modal logic: according to this conception, properties are either functions from possible worlds to possible individuals or simply sets of possible individuals.[6] This conception is comparable to the technical conception of a proposition that I discussed in the last chapter, the one taking them to be functions from possible worlds to truth-values or simply sets of possible worlds. I shall say little about this conception in the present chapter, because it has played almost no part in philosophical debates, either historical or contemporary, about the existence and nature of properties.[7] The remaining conception

---

[1] Kant (1997), B377.
[2] Frege (1892), p. 51.
[3] Russell (1953), p. 200.
[4] Russell took this line in his paper "On the Distinction between Universals and Particulars."
[5] For the first, see Steup (1996), who describes the properties he analyzes as concepts (p. 21), or perhaps Chisholm (1991), p. 169. The recent view espoused by van Inwagen (2004), which is a significant improvement over the view of Steup or Chisholm, also belongs here. For the second sort of view, see Armstrong (1978).
[6] See Lewis (1986), pp. 50-69.
[7] As my subsequent discussion illustrates, I have no objection to this conception when it is used for the technical purposes I mention in the text.

that I shall discuss is a more plausible rival Platonic or Aristotelian conceptions; it take properties to be attribute-instances, or "tropes."

Since each of the three generic conceptions I intend to discuss may be spelled out in a variety of ways sufficiently elaborate to merit the term "theory," I shall discuss them as theories and distinguish them from one another by reference to well-known instances. In deference to Frege, I shall call the first sort of theory an *F-theory*; I could call it a P-theory, after Plato, just as well, but Frege is more representative of the modern theorists I have in mind. In deference both to Aristotle and D. M. Armstrong, I shall call the second sort of theory an *A-theory*. Because the third sort of theory is associated with two quite different names, Donald Williams and Keith Campbell, I shall speak of T-theories, using the letter that begins the word "trope," which is applied to the sort of the entities that such theories postulate. As it happens, A-theories and T-theories appear to be the most popular these days, but I shall contend that a certain kind of F-theory is the best of the lot. In my view A- and T-theories are both undermined by a serious error about predication, which the better F-theories easily avoid. I shall begin with A and T-theories, leaving F-theories to the last.

Armstrong calls the objects of his A-theory immanent universals, but he takes them to be properties or relations.[8] As he understands them, properties are absolutely determinate entities that may exist at many different places at the same time; they are "repeatables." The basic reason he gives for thinking that such things exist is that different particulars have what appears to be the same nature;[9] they are the same in a significant way. This sameness, which "cannot be explained away" in his opinion, may be partial rather than complete, for a red ball and a red book may have something in common too. Normally, a general predicate is applicable to a thing because of some property the thing possesses, but if two things are truly described by certain predicates —"colored," for example—the color-property possessed by one may be very different, he says, from the color-property possessed by the other. One may be green while the other is red.

A T-theory differs from an A-theory in denying that any attribute possessed by one particular is (or could be) identical to an attribute possessed by another particular. For a T-theorist, properties are nonrepeatable entities: each one of them is uniquely instantiated, a unique property-instance. Such instances may be more or less similar, however. If two objects, x and y, are both $\text{scarlet}_{29,}$ the scarlet_{29} of x is an exact duplicate of the scarlet_{29} of y; if x is scarlet_{29} and y is scarlet_{16}, the scarlet of x is very similar to the scarlet of y, but not a duplicate of it.

An F-theory differs from A- and T-theories in denying that properties are literally present in the spatio-temporal world. According to an F-theory, an elementary statement, judgment, or belief "s is P" is true just when the referent of "s" (the subject) *falls under* (or bears some comparable relation to) an F-property that is associated with the predicate "is P." The distinctive feature of an F-property is that it does not exist *in* the particulars that fall under it, exemplify it, or are otherwise related to it. According to some conceptions, F-properties exist in "a realm apart"; according to others, they are items we construct and take account of in deciding whether a predicate is or is not applicable to a particular object. The universals of the F-theory I shall recommend are, in fact, best described by Frege's word "concept," but my use of the word will not be exactly the same as his.

---

[8] Armstrong (1978), vol. 1, p. 6.

[9] "If two things have the very same property, then that property is, in some sense, 'in' each of them," Armstrong (1978), p. 108.

Problems with A-theories and T-theories

A theories and T-theories are similar in locating properties in particulars, but the difficulties they involve are quite different. As I have indicated, A- theories are thought to provide a general explanation of why predicates are truly applicable to particular things. The predicates are applicable because the objects possess appropriate properties as constituents. Some property or other corresponds to every predicate that truly describes the object. This last assertion raises two serious problems, one about particulars and one about the properties they possess.

The problem about particulars can be brought out as follows. If A-properties are actually constituents of particulars, a particular is either a complex of A-properties (as "bundle" of such things) or it contains something in addition to those A-properties. Both alternatives have historically been defended.[10] The first is not plausible by contemporary standards. Leading A-theorists reject it—as Armstrong does[11]--and it is vulnerable to an objection that I shall develop later in connection with T-theories.[12] I shall therefore pass over it now and consider the second alternative—that particulars are something in addition to the A-properties that they possess. The difficulty with this alternative is that it renders particulars unnecessarily mysterious. Particulars become mysterious on this alternative because the nature of a thing, according to A-theories, is constituted by the properties it possesses, but the particular is distinct from those properties. As a result of this, a particular is distinct from its nature--distinct not just in the sense of being not identical with it but in the sense of being something in addition to it. John Locke famously described such distinct particulars as "things I know not what," mere *substrata* that support qualities or provide a subject in which qualities can inhere.[13] He acknowledged that he has no clear and distinct idea of such things, and A-theorists who regard particulars as ultimately "bare" subjects ("bare particulars") describe them in an equally mysterious way. [14]

Armstrong, an A-theorist who accepts the second alternative, thinks that these problematic descriptions can be avoided by distinguishing two conceptions of a particular, one thick and one thin. According to the thick conception, a particular is a "thin" thing along with its qualities: If the thin thing is **a** and **S** is the conjunction of **a**'s qualities, the thick particular is the state of affairs, **a-**having-**S**.[15] According to the thin conception, a particular--in this case, **a**--can be thought of in abstraction from the state of affairs in which it figures; so conceived, it can be thought of as distinct from the properties **S**. Armstrong concedes that, thought of this way, the thing **a** is "perhaps...in a way" a bare particular: "it is the mere thisness of a thing as a Scotist would put it"; it "can have no properties. It is a bare principle of numerical difference."[16] Although Armstrong allows that non-spatio-temporal particulars are imaginable, he nevertheless suggests that the particularity "or thisness" of a particular might in fact be identifiable (owing to the nonexistence of immaterial things) with a "total-position" in space-time. The attributes of such positions, their shape and size, are of course universals, he says; but two different total positions may yet be two, he thinks, even though they have the same attributes.[17]

---

[10] Russell (1940) defended the first alternative; Armstrong (1984), among others, defended the second .
[11] Armstrong (1978), vol. 1, pp. 89-101.
[12] See below, p. 117.
[13] Locke (1984), p. 392.
[14]See Allaire (1963), pp.
[15] Armstrong (1984) p. 254.
[16] *Ibid.*
[17] *Ibid.*

It seems to me that Armstrong's thinly conceived particulars, and therefore the thickly conceived ones of which they are constituents, are every bit as mysterious, ultimately, as Locke's "things I know not what." It is, of course, possible (epistemically speaking) that Armstrong's thin conception of a particular is not really required for a defensible A-theory denying that particulars are complexes of universals. Roderick Chisholm, who spoke of a thing's properties in a way that suggested he held an A-theory himself, said that the following assertions are "simply a muddle":

1. If we distinguish between a thing and its properties, then we must say that the thing is a "bare particular" that doesn't have any properties.
2. One is tempted to regard "This is red" as a subject-predicate proposition, but if one does so, one finds that "this" becomes a substance, an unknown subject in which predicates can inhere....[18]

Chisholm did little to explain why these assertions are muddles other than observing that the idea of a self (a self being a particular) is "the idea of an x such that x loves or hates and such that x feels cold or x feels warm, and so forth."[19] Evidently he was confident that the x he speaks of here is not a bare particular because it is patently not characterless but warm, cold, a lover or hater, and so on. Yet if properties are A-properties, collectively distinguishable from the subject that "has" them, how could Chisholm know that the x he speaks of is not "an unknown subject in which predicates can inhere"?

The claim that something that is warm or cold or wet or dry cannot be a bare particular is perfectly acceptable to me, but then I do not hold an A-theory. As I have explained, those who hold such a theory conceive of properties in a particular way, and they also assume an analysis of predication that makes a mystery of something otherwise not mysterious at all. They take properties to be entities that are "possessed" by particulars but distinguishable from them. When a particular, **a**, is said to be **F**--blue, say--the A-theorist adopting the first alternative interprets the speaker as saying that a property, **u**, is present to **a** but distinguishable not only from it but from the entire "bundle" (or sum) of properties **a** possesses. Although **a** can be known as the possessor of **u** and whatever other properties it may possess, its nature as something distinct from those universals cannot be known because any predicate or concept that one might use to describe its nature is said (by the A-theorist) to refer some other property that is distinct from it or any part of it. So the intrinsic character of **a** remains mysterious, according to the theory.

I said earlier that A-theories also create a problem about properties themselves. The problem, which Leibniz may have been the first to see,[20] can be brought out as follows. According to A-theories, if we are to explain why a general term is truly applicable to a thing, we must ultimately acknowledge the presence in it of some A-property or universal.[21] But A-properties can perform this explanatory role only if they differ from one another: the A- property whose presence in **x** explains

---

[18] Chisholm (1976), pp. 43f. Chisholm says the first argument "seems" to have been offered by Allaire in Allaire (1963); he quotes the second argument from Russell (1948), p. 97.

[19] Chisholm, (1976), p. 39.

[20] According to Mates (1986), Leibniz held that the "accidents" of substances should not themselves be regarded as parts of reality because if we think of them this way we shall have to acknowledge an endless series of higher-order objects, "abstractions of abstractions" (pp. 171-173).

[21] In "To Aune," Armstrong expresses a cautious attitude to this principle, saying "It may be that some such principle is true' (p. 252). But in discussing the view that he calls Predicate Nominalism, he raises the question, "In virtue of what do these general terms apply to the things that they apply to?" implying that a satisfactory answer will have to refer to universals. See Armstrong (1978), vol. 1, p. 19.

why "blue" is applicable to **x** must differ from the A-property whose presence in **y** explains why "red" is applicable to **y.** Similarly, the A-property whose presence in **z** explains why the absolutely determinate predicate "$scarlet_{29}$" (assuming it to be such) applies to it must be the same as the property that explains why this predicate is applicable to some **w** ≠ **z**. But if properties can differ or be identical in this way, they must have features that distinguish them. Since A-theorists assume that things possess features (are thus and so) only if they have appropriate A-properties, such A-properties must be their constituents in just the way that the A-properties of particulars are their constituents. As in the case of particulars, a distinction will have to be drawn between the A-properties and their constituents, and the A-properties will end up with the characterless "thisness" that Armstrong attributes to particulars. Since the constituents comprising the nature of a property must be distinguishable from one another, they too must have different natures, and this means that they will possess constituents in turn. There can be no end to this on A-theorist assumptions: every property will be like an infinitely complex system of Chinese boxes, one within another and each containing its own peculiar "thisness." This consequence is incredible.

Armstrong does not accept this criticism of his theory. When I brought it to his attention in the mid-eighties, he replied that although a fully determinate shade of white, W_{57}, will be different from every other property, the relevant differences may only be "numerical."[22] I find this suggestion unintelligible and certainly at odds with the assumptions about predication implicit in his A-theory. If particulars **x** and **y** could be distinct without having attributes (that is, A-properties) that distinguish them, how could two universals be distinct things without having attributes (that is, A-properties) that distinguish them? A-theorists attribute properties to particulars on general grounds--they want to explain the similarities and differences that are recorded by the application of predicates. We may not have an infinity of predicates that we customarily apply to properties, but that fact is irrelevant to the metaphysical explanation of the similarities and differences that must exist between them if they are to do the explanatory work that A-theorists attribute to them. If $\mathbf{u_1}$ and $\mathbf{u_2}$ are distinct objects with explanatory potential, there must be some **F** that $\mathbf{u_1}$ has but that $\mathbf{u_2}$ does not have--and so on without end.

On the face of it, T-theories (trope theories) do not face the problems I have attributed to A-theories. According to them, particulars are not ultimately mysterious subjects of predication but "bundles" of tropes. But tropes differ from one another in spite of the similarities that may exist among them. They cannot differ or be similar, however, without having definite natures--and this means (given the assumptions of the theory) having distinguishing attributes. If a trope theory is consistent in all its presuppositions, a thing's ostensible attributes are actually its constituents: "a is F" implies that a particular **F-ness** is part of **a**. Consequently, if a T-theory is consistent in this way, it must allow that every trope consists of further tropes--and so on without end. Since unanalyzable particularity can be no more allowable for tropes than for ordinary particulars, every identifiable thing will decompose into a bundle of other things, and no bundle will have an irreducible core. (An analogous consequence will hold for A-theories that regard particulars as complexes of A-properties; this is the objection that applies to "the first alternative" that I did not discuss when I considered A-theories.)

Adopting a defensive strategy similar to one naturally adopted by A-theorists, T-theorists might argue that tropes can resemble and differ without having similar or contrasting components--that their resemblances and differences can be ultimate facts about them. But an exactly parallel argument could be used to argue that or-

[22] Armstrong (1984), p. 252.

dinary particulars can resemble and differ without having tropal constituents: their resemblances and differences can be ultimate facts about them. The latter claim is no less credible than the former. In fact, it is far more credible, all things considered: it does not have the bizarre consequences of a consistently developed trope theory.

Predication

When David Lewis, in his important paper "New Work for a Theory of Universals," criticized Armstrong's main argument for universals, he insisted that predication should be acknowledged as "primitive," as not requiring any analysis, least of all the sort of analysis that Armstrong was tacitly requiring.[23] When you attempt to explain why a thing **a** is **G** by introducing some constituent **u** in **a**, whether A-type or T-type, you are always left with an unexplained datum of the same structure: **u** is **F**. This way of putting the point is closely related to mine; I have simply tried to show what happens when predication is consistently analyzed according to the pattern assumed by an A- or T-theory.[24]

Although I would not attempt to reduce predication to some more basic relation, I don't want to say that I accept it as primitive and let it go at that. The fact that shrewd philosophers constantly provide (or assume) unacceptable analyses of it makes it important to offer some clarification of it--to say enough to help readers resist the tendency to offer a reductive analysis. I also want to say enough to discourage a philosopher from saying, as Armstrong did, that if I say that a dog is barking but "deny the metaphysical reality of properties and relations" I am committed, against my will, to the view that the world consists of "truly bare particulars."[25]

To clarify the basic nature of predication as I understand it, we should consider what is fundamentally accomplished by elementary English sentences having a predicative function. The following examples illustrate the simplest forms that A-theorists make use of in developing their views; they are also employed by T-theorists, but I shall ignore the latter in this context.

(1) Socrates is wise.
(2) Alcibiades laughed.
(3) Plato admired Socrates.

In (1) the predicate contains a linking verb conjoined to an adjective, a construction that A- theorists interpret as relating a subject to a repeatable property or universal. In (2) the predicate is a mere verb, which is less plausibly interpreted by the subject-R-*Fness* paradigm; and in (3) the entire sentence must be transformed to accord smoothly with A-theorist preconceptions: it must assume the form of "The ordered pair <Plato, Socrates> R *admires.*"

---

[23] Lewis (1983). I interpret Lewis' claim that predication should be acknowledged as primitive as equivalent to the assertion that a predication to the effect that a thing *a* is *F* may be incapable of any ontologically more revealing paraphrase. I say "may" rather than "is" because some predications do admit of such paraphrases. "∃x(x is a brother of Tom or x is a sister of Tom)" may be a revealing paraphrase of "Tom is a sibling."

[24] Actually, I made essentially the same point as Lewis in both Aune (1985), p. 44, and in Aune (1984), where I said "Whatever the ultimate entities of the world may be, a proposition of the form "*a* is *F*" must be true of them without implying the existence of further, more elementary entities. If universals did exist, they themselves would be describable by propositions of this form; but the proponent of universals would not insist that such propositions could be true only if entities of a further sort exist. To parody Wittgenstein, "predication has to come to an end somewhere," p. 167.

[25] *Ibid*, p. 254.

In contrast to the A-theorist, I take "is wise," the predicate of (1), to be a unit, one by means of which the person denoted by the subject is described. The predicate does not denote (or pick out) a repeatable component that is attached to this subject; it applies directly to the subject itself, telling us what *the subject* is like. Since a wise person is not a characterless "this" but a wise thing, the predicate of (1) gives no support to the inferences Chisholm regarded as muddles. The same is true of the predicate in (2). Here a simple verb is predicated of a subject: Alcibiades is described as having laughed at some time. If, using the sentence, I describe Alcibiades this way, I cannot reasonably allow that I have described a "bare" particular, for I have described Alcibiades as having laughed, and nothing ultimately characterless can do a thing like that. Sentence (3) is similar to (1) and (2) in describing something, but it describes two people rather than one: it describes Plato in relation to Socrates. It does not identify anything other than Plato and Socrates, and there is no justification for representing its logical structure in the contrived way suggested above.

What I have just said no doubt needs elaboration, for the reasoning supporting the postulation of A-properties is very deeply entrenched in the thought of many philosophers. The key consideration is that the predicates in sentences like (1) and (2) directly apply to the things picked out by their subjects; they do not apply to, stand for, or denote some further items that their subjects may possess. If I say that a fireplug is red, the only thing I am talking about is the fireplug; I am not talking about something that it "has." Anyone who is familiar with red things and understands English will know what I am in effect saying about the plug: *it* is a red thing. Red things resemble one another with respect to color, but one should not suppose that this resemblance consists in a common component, an A-property. The A-theorist Armstrong actually denies that there are generic universals: he claims that repeatable determinate whites (for instance, yellowish white$_{25}$ and greenish white$_{14}$) color-resemble without exemplifying a higher-order whiteness, and a T-theorist would claim that corresponding tropes would color-resemble without containing a common white. I avoid the exotic but make a parallel claim: white things (bedsheets, writing paper) and red things (fire engines, balloons) color-resemble one another without containing any common metaphysical element. If you are familiar with fire engines and can speak English, you will know what I mean in speaking of a red balloon. You will not have to be familiar with any metaphysical entities, particular or general, that supposedly inhere in certain balloons and fire engines.

Armstrong claimed that one cannot avoid postulating A-properties by speaking of color-resemblance or shape-resemblance because these resemblances are merely "respects" in which objects resemble and differ, and such respects require explanation by reference to A-properties. Armstrong's claim is unconvincing, however. When we learn to apply a color vocabulary to the objects around us, we learn to classify *them,* the objects, as more or less similar in color; and we readily learn to classify things as more or less similar in respect to other possible descriptions: for instance, in respect to being round or square. ("Is this as round as that?" we may ask?) What is redder or rounder or squarer than another thing are *particulars*; they are what we are comparing, not some abstract component that they have; and *they* are what resemble and differ in respect of their color or shape, not their supposed abstract components. When we apply predicates, simple or compound, to particulars, we describe *those particulars* (we say what *they* are like). We do the same when we speak of how *they* resemble one another.

It is useless for an A-theorist or T-theorist to reply, "Why do you emphasize that we describe particulars? We don't deny this. We simply insist that particulars are truly described as thus and so because they possess qualities, though we disagree about whether those qualities are repeatable or particular." The reply is use-

less because it assumes that true predication is invariably explained or justified by reference to items other than the particulars that are described.[26] Yet these other items can do the intended work (of explaining or justifying) only if they have natures of their own. If having a nature (or being such and such) is invariably assumed to involve some kind of relation to a higher-order object that must itself have a nature of its own, a single predication is never fully understandable: it must always be understood (or tacitly analyzed) in relation to something further, which must be understood in the same way--and so on without end. If a predication is ever fully understandable--and it usually is--some predication must be understandable in its own terms, without reference to further objects. I contend that "x is round" and "x is scarlet" are acceptable examples of predications understandable this way.

An ostensible reason for postulating A-properties or tropes that I have not mentioned deserves a special comment here. In everyday life we often find it useful to employ singular descriptions that, carelessly considered, seem to apply to something abstract rather than concrete. If we run across a piece of fabric that is colored in a way that is, for us, unusual and especially attractive, we might proceed to use the words "the color of that fabric" in mentioning our discovery to our friends. Although there is nothing mysterious about the fabric that interests us, our talk of "its color" might lead a philosopher to think of G.E. Moore's famous claim that colors are simple, unanalyzable qualities.[27] This claim is quite foreign to what we have in mind when we think about the attractive fabric. We may describe that fabric as red in a very special way: it, the fabric, is a little like this (a crimson lampshade) and a little like that (a little dress that is red but without the orange tinge of something crimson). A philosopher hearing the singular term "the color of the fabric" and thinking of its referent as a object in its own right might be led to Moore's view because our descriptions of the fabric do not equally apply to such an object. The abstract "color" is not a little like the scarlet cloth and a little like the reddish dress. It is evidently not spread out in space, either, nor does it reflect light and look a bit different when the fabric is moved the about the room to see how it contrasts with objects that are blue or yellow. The object the philosopher thinks of is really nothing like the fabric that we found so special on account of *its* color. The singular term we used is very misleading to the literal-minded philosopher.

Advantages of F-theories

The critical remarks I have been making in the preceding pages do not apply to F-theories of properties and relations, the sort of entity Frege described by the word "Begriff" or (as we would say) "concept." Instead of postulating items whose presence in a thing are supposed to account for the truth of what is said about that thing, an F theory seems to be built on the idea that what accounts for such a truth is the thing itself.[28] Objects are truly described as blue, red, or green because *they* are blue, red, or green—not because they possess some further items that possesses some other feature. It is, of course, true that all objects truly described as blue fall

---

[26] This claim would not be made by my colleague Jonathan Shaffer, who tells me he postulates tropes only to account for the causal properties (or interactions) of empirical objects: he does not suppose that a trope corresponds to every true predication. In opposition to his view, I say that the special tropes he recognizes are excess baggage, for a thing's causal interactions are adequately explainable by reference to its own empirical character: a window shatters, for example, because *it* is brittle and struck by a sufficiently heavy object. The same principles apply to the interactions of micro-entities: they behave as they do because of what *they* are like. No special tropes are needed.

[27] Moore (1903), p. 7.

[28] There are, of course, certain cases in which a thing satisfies the conditions for being **F** only if it is related to a further thing: to be a brother one must be suitably related to another person. The point is simply that there is no general requirement to this effect.

under the concept *blue*, but saying that they fall under this concept just means that the concept is applicable to them. It applies to them because they are the right sort of object: the concept *blue* applies to blue things.

Another virtue of F theories, at least the kind espoused by Frege, is that the "concepts" they postulate are plausible constituents of propositions.[29] As I noted in the last chapter, the singular proposition that Kaplan takes to correspond to "Socrates is wise" is an ordered couple whose first member is Socrates and whose second member is the property *wise.* But if this property were the sort of thing postulated by an A-theorist—namely, the sort of thing whose presence in an object makes it the sort of object it is—we would expect the couple to be a wise one, just as we would expect an object containing the A-property *red* to be a red one. Of course, an A-theorist would want to dash these expectations. But if a property were the sort of thing he says it is, we would need an illuminating explanation of why our expectations are erroneous. Just saying that a property is not present in a proposition in the way that it is present in a particular is hardly sufficient.[30] If the A-property *red* is a red-making property, as A-theorists claim, it should have that character wherever it is located—whether in an apple or a proposition.

Although a Fregean concept is not a psychological entity, it is like a Kantian concept in relating a predicate to the class of objects of which it is true or to which it applies. As Frege is commonly interpreted,[31] the sense of a sentence (its meaning, the thought it conveys) is a proposition, and a concept, being the sense of a predicate, is an appropriate ingredient in a proposition, one that helps determine the sentence's reference. How does the concept do this? It does it by identifying one of the constituents by means of which the sentence's reference is identified. For Frege, this reference is a truth-value. Here is an example of how such a truth-value is identified. Take the sentence "Socrates is a man." The subject of this sentence, the name "Socrates," has an individual concept as its sense; this individual concept picks out the man Socrates as the referent of that name. The predicate, "is a man," has the property *man* as its sense; this sense (this concept) picks out the class of men falling under this predicate. (Technically, the class is the reference of the predicate.) By virtue of the way the concepts are connected in the proposition, the reference of the sentence is determined to be truth or "the true" (to have this value) just in case the reference of "Socrates" is a member of the class that is the reference of the predicate. Since Socrates is a member of this class, the sentence is true.

In spite of being similar to a Kantian concept in a central respect, a Fregean concept is different in an important way. According to Kant, general ideas (or concepts) are distinguished from singular ideas (intuitions in his sense) in being related to objects only "mediately," by means of "marks" that are common to those objects.[32] Kant's marks recall A-properties, which have no place in Frege's system. For Frege, the objects to which a predicate applies are simply the objects falling under the concept that is the predicate's sense. As I have explained, Frege identified a thing's properties with the concepts under which it falls. As he put it, "to be Φ is a

---

[29] The properties espoused by van Inwagen (2004) are precisely this: constituents of propositions, which he takes to be "things that can be said." I comment on propositions understood this way in Appendix 4.

[30] To accommodate his supposed direct awareness of the incompatibility of red and green, BonJour (1998) tentatively suggests that a thought whose content affirms this incompatibility "instantiates a complex universal of which redness and greenness are literal constituents." He emphasizes that the colors are exemplified in a special way, so that they do not make his thought literally red and green. But he does not explain how this special exemplification is accomplished, or how greenness can be present, and recognized to be present, in thought or consciousness, where nothing is actually green (p. 185).

[31] The common interpretation I refer to is the one given by Carnap and Church; see footnote 45 of chapter three above.

[32] Kant (1997), A320, B377.

property of Γ" is just another way of saying "Γ falls under the concept of a Φ."[33] Frege used Kant's terminology of conceptual "marks" in his analysis, but he regarded marks as components of compound concepts, not items shared by the objects falling under a concept.[34]

As these observations indicate, a basic function of concepts in Frege's system is to connect predicates with the objects to which they apply. How do they do this? Frege gave only a partial answer. He said (according to the Carnap-Church interpretation) that a predicate has a concept as its sense, but he offered no explanation of how a concept is connected to the objects falling under it. The notion of an object falling under a concept was primitive in his system, an unexplained explainer. He was no doubt convinced that we learn to identify what falls under a concept in the process of learning to understand it. If a concept is not complex, we learn to apply it to instances directly—as we learn to apply the concept *red* to red objects. This does not of course explain why an object **o** falls under concept **A** rather than concept **B**. It gives us no understanding of the connection.

Did Kant explain this connection better with his twin notions of concepts and marks? The answer is no. He evidently believed that empirical concepts are associated with specific mediating marks because, according to the erroneous theory of concept-formation that he accepted for empirical concepts, these marks provide the experiential material from which the appropriate concepts were created by a process of abstraction.[35] The redness perceived in roses and sunsets is the sensory input, Kant thought, from which the concept of red is ultimately created. But input or history does not explain reference, any more than similarity does.[36] If we want an explanation of how elementary predicates become attached to these rather than those objects, we must go beyond Kant and Frege and construct such an explanation ourselves.

It is worth noting that the role of concepts in recent possible-world semantics leaves the connection between predicates and their extension just as unexplained as Frege did. Just as propositions are identified, by these theories, with functions from worlds to truth-values (or simply as sets of worlds, the ones in or at which particular sentences are true), so concepts, understood as the senses or intensions of predicates, are identified with functions from worlds to sets of individuals, the individuals to which particular predicates apply in each world.[37] The concept *green*, or the intension of the predicate "is green," associates with a world **w** the class of objects that are green at that world. The function here is essentially a correlator, if it is not a mere class of correlated items, and it provides no explanatory account of how this or that object in this or that world is ultimately attached to this or that predicate. To say this is not to expose a defect in these semantic theories, for they make no pretense of providing such an explanation. But an appropriate explanation is important for epistemology.

What are Concepts?

To develop the desired explanation I want to begin with Frege's notion of a concept and, by means of various qualifications and explanations, work my way toward a conception of my own. In current philosophical practice the word "concept" is used very loosely and equivocally, applying to ideas, abstract objects of conception,

---

[33] Frege (1892), p. 51.
[34] Ibid.
[35] See Kant (1974), pp. 99. For a criticism of abstractionism, see Geach (1956), chs. 6-11.
[36] See Putnam (1981), ch. 1.
[37] Or simply as sets of possible individuals, as Lewis (1986) says, pp. 50-69.

and sometimes even uses of words.[38] In spite of this ambiguity and indefiniteness, it nevertheless has connotations that I want my notion to preserve. Judging from an observation by Elizabeth Anscombe, the terminology of objects falling under concepts is not unusual in everyday German. She reported that Michael Dummett once saw in a Münster railway station a notice beginning "All objects that fall under the concept *hand-luggage*...." (*Alle Gegenstände, die unter den Bergriff Handgepäck fallen...).*[39] This anecdote reminds us that we commonly classify things by "concepts" that are humanly invented and rest on conventions having significance only for special groups. Frege no doubt thought of concepts this way even though his technical notion of a *Begriff* was supposed to be a special "logical" one, not identical with a vernacular idea.[40]

The word *hand-luggage*[41] gives expression to a humanly-invented, nontechnical concept, and so does *carry-on*, *personal effects*, *engagement ring*, *American citizen*, *slave-driver*, *mule*, *Venetian blind*, *flotsam*, *jetsam*, *retriever*, *pointer*, *barber*, *typewriter*, *zipper, computer, computer-programmer*, *disk jockey,* and *playboy*--this list can be extended as far as you want. It is true that things in nature fall under these concepts, but it is absurd to suppose that these concepts are eternal entities that define the structure of reality, as Plato's Forms were supposed to do. They all have histories, and they came into existence as the result of numerous contingencies.

Another important feature of the everyday words we use to classify objects is that they are vague. A vague word, as I observed earlier, is one that clearly applies to some actual or imaginable things, that clearly fails to apply to some such things, and that neither clearly applies nor clearly fails to apply to other such things.[42] *Bald* and *tall* are standard examples of vague words, but in fact every generic color word is vague, and so are most of the words we use in everyday life. Consider such words as *sarcastic, sardonic, frivolous, trivial, flimsy, superficial, paltry, petty, trifling, lucky, unimportant, yuppie, dismal, morose, severe, zaney, dour, carefree, windy, brisk, sparse*--again, the list can be continued almost endlessly. Although such words can perfectly well express vague concepts or ideas, they cannot stand for definite properties or items in reality, because they connote nothing definite or determinate.

One might suppose that a word like *bald* could be construed to apply in a strict sense to people whose head is utterly hairless and to apply to people with some hair only in a loose and popular sense. But the word is not really used this way; and an analogous claim holds for vague words generally. Take the word "sarcastic," which is familiar to every adult speaker of English and is used with confidence even by high school students. No adult or adolescent has any doubt about its application to some people and some things people say. Some people are clearly sarcastic either generally or on some occasions; some people clearly are not; and a great many people exhibit behavior that is not clearly classifiable either way. In spite of the confidence with which "sarcastic" is commonly used, it is a very difficult word to define or even clarify by synonyms. Its etymology is very illuminating, however. It derives from the Greek *sarcazô*, which Lidell and Scott define as "to rend of flesh" in the manner of dogs.[43] As this derivation indicates, "sarcastic" was originally

---

[38] See Margolis and Laurence (1999).
[39] Anscombe (1959), p. 122.
[40] Frege (1982), p. 42
[41] I normally mention words by putting them in quotation marks, but I mention so many words in this section that I use the alternative method of italicization here.
[42] See chapter one, footnote 14.
[43] Liddell and Scott (1984), p. 630.

metaphorical. The metaphor is very tenuous today, but we still think of a sarcastic remark as one that is wounding, hurtful (and a sarcastic person as someone prone to making such remarks). Since we have encountered many clear cases of sarcasm and non-sarcasm, we have the ability to recognize such cases when we see them; but we are constantly presented with borderline cases that we cannot confidently classify either way. *Most of our vocabulary is like this*. Our words commonly involve metaphors--compare *inspire, inspiration; expire, expiration; understand, understanding*-- and their meaning is rarely precise or determinate.

To be confident that the kind of F-theory I wish to recommend is actually a good one, I must be more specific about what I take a concept to be. As I noted, there is no definite and unequivocal sense in which the word is normally used in philosophy. Insofar as a concept is assumed to be something that a person may possess, there is general agreement that the relevant mental capacity is associated with general words. A person who understands the adjective "red" is said to have the concept of red, and this same concept is said to be possessed by someone who understands a word synonymous with "red." If we accept this presumption, we can say that the concept *red* is something associated with "red" and its counterparts in other languages. The question is, "What is the 'something' and how is it associated with the relevant words?"

One way of answering the question is suggested by the observation that a person who uses the word "red" in speaking or thinking would generally be held to be *employing* the concept *red*. A French person who uses "rouge" would be said to employ the same concept. Now, if "rouge" is a good translation of "red," the words are used in formally analogous ways. Speakers of French apply "rouge" to objects that speakers of English would describe as red, and each would relate their word to other words of their language in a way that is parallel, formally speaking, to the usage of the other. Thus, the French would use "rouge" in relation to "vert" in basically the way that we use "red" in relation to "green." It is convenient to have a general term by which to classify words that are functional counterparts in this way. Such a term was supplied years ago by Wilfrid Sellars; he constructed it by means of his dot quotes: any expression that is a functional counterpart to "red" can be described as a *red*.[44] (I use asterisks where Sellars uses dot-quotes, because asterisks are easier to see.) If we use Sellars' terminology, we can say that the concept *red* is something that is closely associated with the use of *red*s.

D. M. Armstrong once said that the task of giving an account of "the" type-token distinction is a "compulsory question on the [philosopher's] examination paper."[45] A plausible way of relating *red*s to the concept *red* is to say that the latter is the type of which the former are tokens. Saying this requires that one come to terms with *a* type-token distinction (there may be more than one), but it accords with the common assumption that if you understand and use "red," you have and employ the concept *red*, and that if you understand and employ "rouge," you have and employ the same concept.

When we think of types, we often describe them in ways appropriate to tokens. This tendency is perhaps evident in Plato's practice of describing particulars as imperfect imitations of perfect Forms,[46] but it stands out in Hilaire Belloc's amusing lines:

> The llama is a woolly sort of fleecy hairy goat,
> With an indolent expression and an undulating throat.[47]

---

[44] See Sellars (1979), ch. 4.
[45] Armstrong (1978), vol. 1, p. 17.
[46] As in *Republic* 597a.
[47] "The Llama," in Belloc (1970), p. 245.

It is obvious that what is said of the type here is properly predicated of the tokens, for only particular llamas are fleecy hairy goats with indolent expressions and undulating throats. Surely no abstract object is hairy and has an indolent expression. Wilfrid Sellars devoted a lot of attention to expressions such as "the llama"; he called them *distributive singular terms* (or *DSTs*) and said that statements containing them are by definition equivalent to statements about concrete things. In his view a statement of the form "The llama is F" can be paraphrased as "Llamas are F."[48]. This view is very plausible, I think, for the predicate in "The llama is a woolly sort of fleecy hairy goat" certainly makes it appear that the statement as a whole is about actual llamas. If "the concept *red"* can also be understood as a DST, statements containing it can reasonably be regarded as equivalent to general statements about the expressions said to "express" that concept. This will fit in nicely with the conceptualist view of propositional content that I presented near the end of the last chapter.

As attractive as I find this suggestion about concepts to be, I must acknowledge that the distributive treatment Sellars and others[49] have endorsed for words ostensibly referring to types has been seriously questioned in the literature. Linda Wetzel has in fact criticized it at length in an earlier article[50] and in a more recent book.[51] Before attempting to develop the suggestion about concepts, I must obviously come to terms with this criticism.

Some Problems about DSTs

According to Wetzel, the schema commonly offered for the elimination of DSTs is seriously defective. It fails, she says, because the relevant tokens do not always possess the attributes ostensibly attributed to the type. To take the example from Hilaire Belloc, although it is perhaps true that

> The llama is a wooly sort of fleecy hairy goat
> With an indolent expression and an undulating throat,

it is certainly not true that every actual llama satisfies this description. Shaved or burned llamas are not wooly and fleecy; beaten llamas do not have indolent expressions; and starved ones probably lack undulating throats. This criticism of the standard definition schema is clearly correct. Even in cases where the ingredient general term seems to apply to all members of a class, the relevant class appears to be restricted to typical or ideal examples.[52] If such favored llamas have wooly, fleecy coats, we can say that "the" llama has such a coat; if we are justified in making this last assertion, we can justifiably conclude that all favored llamas have such a coat. Our "the" statement thus has the assertive content of a universal statement restricted to a domain of favored cases.[53]

---

[48] Sellars (1979), 89-99.

[49] Goodman (1951), pp. 360-63, makes use of essentially the same idea in his treatment of assertions ostensibly about statements.

[50] Wetzel (2000).

[51] Wetzel (2003).

[52] I argued this in Aune (2002). Frege noted it much earlier; see Frege (1892), where he said, "The horse is a four-legged animal" is "probably best regarded" as expressing a universal judgment, say "All properly constituted horses are four-legged animals" (p. 45).

[53] Wetzel shows that "The K is F" does not imply "all Ks are F"; the implication evidently does not go the other way either, since some things truly predicable of every grizzly are apparently not predicable of "the" grizzly. Contingent, accidental features seem to be exceptions. If every actual grizzly happened to lose a claw in a trap or a fight, I doubt we would say that the grizzly lacks a claw.

Wetzel also has objections to this qualified view, however. Her first objection is that the notion of what is normal or properly constituted--and therefore what is ideal-- should be viewed with suspicion; it is not, she suggests, scientifically credible (p. 98). She might be right that these notions are scientifically dubious, but the corresponding distributive statements (the ones about *the* llama or *the* grizzly) would evidently be scientifically dubious as well. If I say the llama has an indolent expression and an undulating throat, a hectoring critic might say, "Okay, Aune, how do you identify a *typical* llama, or a "good example" of the species?" Since I am not an expert on domestic animals, I would have to appeal to someone who is. But I don't think even an expert can provide a definition than can single out typical, or "good," llamas with precision. I say this because I think the notion of a typical or good instance of something is vague, and I expect that even llama breeders might disagree about the qualities llamas should ideally have--just as Airedale breeders do disagree about the qualities Airedales should ideally have, some thinking that, because they are terriers, Airedales ought not to be the eighty to ninety pound animals that others admire. Belloc's statement about "the" llama, like ordinary statements about the cat or the Airedale, is not strict or precise. It calls attention, in an amusing way, to striking features of the healthy, well-cared-for llamas that one might see in a field or a zoo--but it does not pretend to be scientifically exact.

The imprecision of ostensible type statements sometimes leads to problems about verification. In her discussion of statements about *the* grizzly*,* Wetzel emphasizes that not that all grizzlies are big, not all are brown, and not all have humps. Yet it is still true, she insists, that *the* grizzly is a big, humped brown bear native to North America (p. 96). But how do we *know* that this is true? Are we not generalizing from some grizzlies or other? In this case I should say yes, though in the case of the llama and the Airedale, which have been bred to suit human purposes, our conception of "the" animal is partly based on our wants rather than our observations. But there are often striking differences between the instances--the good examples -- from which we generalize. Some relevant differences are associated with sex. When we think of a Black Widow spider, for instance, we are probably thinking of the female, for the males are small, insignificant, and eaten by the female at the completion of the inseminating act. Yet the Black Widow species contains males as well as females. I suspect that we simply ignore sex (we abstract from it) when we make statements about *the* Black Widow spider. When sexual differences are brought up, we are apt to make more restrictive statements. We would probably do the same if we discovered that most female grizzlies do not have humps. Instead of speaking about "the" grizzly generally, we might then speak about the male grizzly, the female grizzly, and possibly even the adolescent grizzly, the cub grizzly, and the aged grizzly (male or female)--if there are distinctive traits that such grizzlies possess.

This brings me to another of Wetzel's objections to the distributive analysis. She says, in effect, that such analyses fail because some properties of the type are derived from the *distribution* rather than the *common features* of its tokens. To support her claim she says that *Ursus horribilis*, the grizzly bear, "had at one time a U.S. range of most of the West, and numbered 10,000 in California alone. Today its range is Montana, Wyoming, and Idaho, and numbers less than 1000. [But no] …particular flesh and blood bear numbers 1,000 or had a range comprising most of the West" (p. 102). Her example here is convincing if her opponents are expected to apply a distributional analysis in a mechanical way, but if they are allowed to use their ingenuity in interpreting predicates, a broadly distributional reading is easily achieved. Take the assertion "The grizzly bear once ranged over most of the western U.S." Put in vernacular terms, this tells us that grizzly bears once ranged over most of the western U.S. Saying that *they* had this range is not saying that each one had this range; the predicate is applicable to the grizzlies collectively: *they* were

distributed over this area. The predicate of the second statement is also collective, a plural predicate taking a plural subject: *they* (certain grizzlies) numbered 10,000 in California alone. The same principles apply to the two statements about the grizzly today: grizzlies now have three states as their range, and they now number 1,000. These collective predications are, of course, reducible to singular ones. Saying that grizzlies *are* distributed over a certain area amounts to saying that individual grizzlies exist here and there throughout that range.

Reflection convinces me that not all statements about "the" grizzly are distributional in the ways I have so far described. If one says that the grizzly was seen in Washington State in 1975, one is not saying that typical instances were seen there then; one is saying that *some* instance was so seen. And if one says (as I heard someone recently say) "Many rich people now transfer nothing to the poor," one is evidently speaking of the poor collectively rather than individually, although one is certainly implying that no poor person is receiving any goods or money from certain rich people. As I see it, there is considerable ambiguity to terms like "the poor" and "the grizzly," and no single distributive analysis is applicable to all of them.[54]

Wetzel agrees that many assertions ostensibly about types can be paraphrased by assertions about tokens, but she insists that we can have no assurance that this can always be done unless we have a systematic way of doing so. As I have implied in my last paragraph, I do not believe that a systematic way of providing such paraphrases can be found; but I have no doubt that the predicates included in Wetzel's favored examples of ostensible type terms apply only to particulars, to "tokens." Only individual grizzlies can be found in the United States (only they can have such a range) and only they can scratch, bite, and become more or less numerous. If the relevant "the" statements cannot be interpreted as saying something about tokens, they will not make sense and they cannot be true. The lack of a systematic means of paraphrasing all examples will not, therefore, at least as I see it, support a commitment to irreducible types. The requirement of a systematic paraphrase for everyday assertions ostensibly about types is, in any case, excessively demanding.

More about Concepts

Before discussing the problems Wetzel raised with distributional analyses of statements containing terms such as "the llama" or "the grizzly," I offered the suggestion that what is ostensibly true of concepts reduces to what is true of certain tokens, specifically certain general terms. The idea seems reasonable in view of some standard assumptions about concepts. Jacques has the concept *snow* when and only when he understands some general term, perhaps "neige," that is a *snow*; Jacques and Tom have a common concept when and only when they understand general terms that are functional counterparts; and I have a concept that is applicable to snow when and only when I have a general term that is applicable to it.

Although the idea that I just mentioned may seem reasonable, it is actually over-simplified in two important ways. The first complication is that concepts have a kind of generality that is greater than anything narrowly verbal. Sellars, who used dot quotes to create special predicates applicable to tokens that are functional counterparts, included mental tokens as well as physical ones. He did this because he was convinced that we can think what we can say and that we can do so without saying anything to ourselves in the way we mentally say something when we silently

[54] It is possible to "reconstruct" statements containing such terms in a way that justifies a distributional analysis; Jeffrey Sicha has done this in an unpublished letter to me. Sicha's reconstruction diverges from everyday discourse in certain respects, as reconstructions commonly do, but I can find no serious fault with it.

recite a poem. If, without uttering anything, we think, "That snow is yellow," we are employing concepts of snow and of something yellow, and doing this requires that certain elements of our thought do the functional work of the words "snow" and "yellow." These elements are reasonably described as *snows*s and *yellow*s, Sellars thought, even though they differ from audible *snows*s and *yellow*s in material (that is, nonfunctional) respects.

I think Sellars was clearly right on this matter: we can think exactly what we can say, and we can do this without saying something to ourselves in the way we mentally say something when we recite a poem to ourselves.[55] Endorsing Sellars's view of the components of a complete thought amounts to asserting that the concepts exercised in thinking and saying the same thing are the same. I accept this view, and I draw the obvious conclusion: What falls under an idea or thought-component in such a thought also falls under a word or phrase in the corresponding utterance. Another way of putting this is to say that the ideas or thought-components involved in verbally expressible thoughts have the semantical properties of the words in the statements that are said to "express" them. They are about the same things; they have corresponding implications; and they are equally appropriate, semantically, in particular circumstances. (The thought "That's red" is just as inappropriate to a green object as the statement "That's red.") As a result of these semantic parallels, the extended view of analyticity I presented at the end of the last chapter applies to thinking in basically the way it applies to statements. No alternative account is needed.

The other complication in the relation of concepts to corresponding tokens is that the latter need not be understood as mental or verbal elements having the semantic character of discrete general terms. Although many of the concepts I have are associated with single general terms, my conceptual resources greatly outstrip my finite vocabulary. The point stands out if we think of concepts as F-properties. A thing has an F-property just when it is truly describable in a certain way, and to be so describable is to fall under appropriate concepts. But the languages in which a thing is truly describable in a way W need not possess discrete predicates associated with a W way of being; a thing can be truly described by a string of words newly invented for the purpose. Strings of words newly put together in the right arrangements can often serve as *definienda* for new predicates, but until those new predicates are introduced, the conceptual specifications will exist only in a long-winded form. We will have the concepts without a simple way to express them.

Since I reintroduced the notion of an F-property, I might just as well comment on a question often addressed to defenders of F-properties. If a theory of this kind is correct, must we say that things would lack all properties if languages or thinking beings did not exist anywhere in the universe? The answer is yes, but it is not a source of concern. A world without F-properties would not be a strange or peculiar world. Each thing existing in it would be intrinsically the same as it would be in a world with speakers and languages. A frozen lake would be just as frozen and just as blue and cold as it now is if the concepts **frozen, cold,** and **blue** had never been invented. The existence of languages does not change anything nonlinguistic. Lakes are what they are independently of how anyone describes them.[56]

Concepts, Predicates, and the World

When I discussed the role of general concepts in Frege's semantical theory, I said that a basic function they perform there is to connect a predicate with the ob-

[55] This same view appears to be held by Jerry Fodor: see Fodor (2000).
[56] I defend this point at length in Aune (1985), pp. 126-130.

jects to which it applies. As for how a general concept can do this, I added that Frege had nothing to say: the idea of an object falling under a concept is primitive in his system, an unexplained explainer. Yet if his concepts are viewed in the way I am recommending, as distributive objects, the idea that they are what relate words to objects is a mistaken one. The truth is, in fact, the opposite: objects fall under general concepts only in the sense that they are truly described or rightly classified by appropriate predicates or phrases containing predicates. Predicates, as I see them, are *direct* describers or *direct* classifiers of objects: they describe or classify without the help of conceptual mediators. In this respect they are semantically comparable to names and demonstratives. Just as the latter directly refer to their bearers, so predicates directly describe or directly classify the objects to which they apply. In neither case is the semantic function performed by means of conceptual intermediaries.

But how, one may ask, does a predicate gain the predicative function it possesses? My answer is that it gets it from they way it is used by speakers and writers. When I criticized the rationalist account of how color-incompatibilities are known, I introduced a simple example that shows how classifying behavior may fix distinctive meanings and create Fregean properties. The example concerned Mary, Tom, and Harry, who described the color of a bush in three conflicting ways. Mary insisted the color was a shade of yellow, Tom a shade of green, Harry a shade consisting of both yellow and green. In stubbornly holding to these contrasting descriptions, Mary, Tom, and Harry in effect distinguished four generic color concepts. Mary's concept of yellow applied to the shrub's unusual shade; her concept of green did not apply to it. Tom's concepts of yellow and green applied in an opposite way: his green included it and his yellow excluded it. Harry, claiming that the shade involved green and yellow in equal degrees, showed that his concepts of yellow and green overlapped in this instance. In holding to their descriptions, Mary, Tom, and Harry jointly distinguished six generic colors and therefore six generic color concepts. These six concepts, we may assume, are closely related, because their extensions (the things to which they apply) overlap in most cases. But they differ in their application to the bush, and this is enough to distinguish them. *Strictly speaking*, they are different.

When I introduced the example, I assumed that the existing uses of "yellow" and "green" are insufficiently determinate to render any one of the three descriptions right or wrong. In holding fast to their descriptions, the three persons were making their concepts of yellow and green determinate in different ways. But even in cases where "correct" and "incorrect" may be thought to apply, the standards often differ for different speakers or writers. Bookish people, those who frequently consult dictionaries, are strongly influenced by etymology and precedent; they are apt to emphasize the differences that distinguish the meanings of such closely related modifiers as "accidental, "inadvertent," and "by mistake."[57] Others, particularly those who rarely consult dictionaries or, as Kingsley Amis once complained,[58] may not even own them, are apt to regard these expressions as basically synonymous: they and their intimates may use them almost interchangeably. The bookish people will say that this indiscriminate usage is incorrect, that those who speak this way are misusing the English language. People who write dictionaries these days have a much more democratic attitude. Usage rules. Those who wish to respect etymology and follow the precedent of fastidious writers may continue to do so, and like-minded readers or hearers will appreciate their fastidiousness. But those who are attuned to a more

[57] See "A Plea for Excuses" in Austin (1961).

[58] Kingsley Amis (1998) complained that "the habit of owning and even consulting a dictionary has largely died out among the general public" (p. 47).

popular idiom have their own standards, and they often find the usage of the others archaic, unnatural, and puzzling.

Meaning, Intending, and Content Clauses

My claim that predicates, like singular terms, directly apply to objects in the world fits together neatly with the conceptualist view of propositional attitudes that I endorsed in the last chapter.[59] According to that view, mental states such as believing, judging, or supposing have "contents" rather than "objects." Philosophers who speak of contents this way commonly say they are given (or conveyed) by declarative clauses attached to a noun clause denoting a person's propositional attitude. They are also given by such a clause in a sentence that may be used to ascribe a propositional attitude to a person—for instance, by "Tom believes that snow is white." When I introduced the conceptualist view, I said that the content of a person's belief had concepts as constituents, meaning by concepts the sort of singular and general ideas that Kant called "Begriffe." This way of describing the contents of a propositional attitude is actually oversimplified in important ways, and it is now appropriate to make some of the requisite qualifications. The qualifications turn out to be epistemically important.

The sentences used to specify the contents of psychological states are structurally the same as sentences used to specify the contents of verbal acts such as stating, asserting, declaring, or, more simply, just saying something. As it happens, the qualifications I want to make about psychological contents are easiest to identify by means of this latter kind of sentence. So consider the following:

Mary said, "Tom has a perverse sense of humor."

According to the terminology of traditional grammar, this sentence purports to contain a "direct" quotation of Mary's words on a certain occasion. The quotation is called an *oratio recta* clause in older textbooks.[60] It is used to describe what Mary is supposed to have said, and it does this by exhibiting the words she is thought to have used in her speech act. If we understand those words, we can understand her utterance and thus know what she is supposed to have said. Now consider the related construction:

Mary said that Tom has a perverse sense of humor.

According to tradition, this sentence contains an "indirect" quotation of Mary's supposed words. This indirect quotation, called an *oratio obliqua* clause in the older books, does not purport to contain Mary's actual words, but it may contain some of them, or a translation of some of them, together with modifications that reflect her relation to the person reporting her statement. If she used French in speaking of Tom, the *oratio obliqua* clause might consist mostly of translated words; if she spoke in English, the words might be more or less the same as Tom's.

A different example can show how much the *oratio obliqua* clause might differ from the original words even when the narrator and the original speaker share the same language. Suppose Betty, pointing to a streetcorner some distance away, told Sam on Monday, "I will meet you there at noon tomorrow." On Tuesday Sam appears on the designated corner at noon, but Betty does not appear. Sam waits for a while, then leaves, and returns an hour later. Betty is there and asks, "Where have you been?" Sam replies, "Where have you been? You told me that you would met me here at noon, but you failed to appear." "I am sorry," Betty replies; "I thought I said one o'clock." This little exchange is perfectly natural, linguistically. Betty does not object to Sam's description of what she told him, but the words he used to report

[59] See chapter 3, pp. 121-124.

[60] Smyth (1920) discusses both *oratio recta* and *oratio obliqua* clauses on p. 584.

her statement agree with her words mostly in the things they identify; only three of them were the same as words she used: "meet," "at", and "noon."

A slightly more complicated example gives a better idea of the differences that are commonly tolerated in indirect quotations of a speaker's statement. Imagine that Fred once said to Sally:

"I will meet you tomorrow at The Blue Moon Café."

Sally, on the following day wishing to convey what Fred told her but forgetting the name of the café he spoke of, might say:

"Fred said he would meet me today at that shabby café on 14th Street."

Although Sally's description of what Fred said would be considered acceptable if her hearer knew what café she was referring to by "that shabby café on 14th Street" and did not dispute the aptness of the description, the only words of Fred's that she repeats here are "meet" and "at," a total of two out of eleven—one less than the total repeated in the simpler example featuring Betty and Sam. To identify the day Fred was speaking of a day after he spoke of it, Sally used "today" in place of his "tomorrow," and she also used pronouns and verbs appropriate to her perspective on him and what he spoke of in place of the ones he used. These differences are just as striking as the differing descriptions of the café on 14th Street. In spite of them—in fact, because of them—Sally's *oratio obliqua* clause gives her hearers a good sense of what Fred was referring to and what he wanted Sally to understand when he spoke as he did. This is what giving the content of an utterance commonly amounts to.[61]

The propositional clauses that give the content of a person's thoughts or beliefs can be understood in an analogous way. If you believe that snow is white, you need not have the thought of snow constantly in mind; but if you are in some way prompted to think about snow or typical examples of white things, the thought that snow is white is apt to cross your mind and serve as a premise in the reasoning by means of which your belief results in overt behavior. The content ascribed to your belief is owing to the content of this distinctive thought, for it, not the dispositional state of believing itself, has the conceptual structure approximately exemplified in the *oratio obliqua* clause conveying that content.[62] The constituents of this conceptual structure are related to the words of the *oratio obliqua* clause in a way that is formally the same as the way the words of an original statement are related to the words of the clause "indirectly quoting" it. Constituent ideas have the same reference as corresponding words, and they frequently possess corresponding implications. These semantic parallels are generally not tight; they can possess all the differences that distinguish Fred's "I will meet you tomorrow at the Blue Moon Café" from Mary's "…he would meet me today at that shabby café on 14th Street."[63]

These observations about the structure and function of *oratio obliqua* clauses have interesting implications for a conception of propositions akin to the conception of concepts that I developed earlier. Like my conception of concepts, which was indebted to the semantical views of Wilfrid Sellars, the relevant conception of proposi-

---

[61] After I wrote the above, I discovered Joseph Almog's wonderful paper, "Is a Unified Description of Language-and-Thought Possible? (2005). Although Almog approaches the subject of indirect discourse from a perspective somewhat different from mine, we arrive at essentially the same conclusion. Almog says "The reporter gets things right as long as he, with *his* terms—[his] names or descriptions—keeps the same denotation as those of the thinker [or speaker]," p. 525. I believe this is true as a general matter, but there are exceptions. If the speaker would reject a denotation, a reporter should not use it. Think of John Perry's example of the essential indexical. See Perry (1979).

[62] I defend this view of believing in Aune (1990).

[63] Almog emphasizes these differences by an example in which he imagines himself saying to David Kaplan after returning from a trip to Paris, "Pierre believes that you are here." Pierre had said to him, he supposes, "Le professeur qui nous a parlait la semaine dernière avec ce costume elégant de Saville Row est retour dans le cité des anges," p. 506.

tions has the same source.[64] It takes propositions to be distributive objects. Instead of implying that propositional attitudes with the same "content" literally possess a distinctive common constituent, the distributive conception implies that these attitudes are tokens of a distributive type. The English-speaker's assertion "Snow is white" is a *Snow is white*, and so is the French-speaker's "La neige est blanche": the statements are semantic counterparts and thus are classifiable by the same illustrating common noun. The propositions in question are distributive objects because what is ostensibly predicated of them is ultimately reducible to predications about corresponding attitudes. The singular term "the proposition that snow is white" is construed (reconstructed) as meaning "the [proposition] *snow is white*, and what is predicated of the latter is taken to be reducible to assertions about appropriate tokens. As an illustration of this, the sentence "The proposition that snow is white is about snow" is reducible to "All *snow is white*s are about snow."

If natural languages did not contain demonstratives or tensed verbs, and if acceptable *oratio obliqua* clauses were so tightly related to the words or thoughts whose supposed "content" they report that a reference to the Blue Moon Café could not be reported by "that shabby café on 14th Street," I would find Sellars's conception of propositions fully acceptable. But natural languages do contain demonstratives and tensed verbs, and acceptable *oratio obliqua* clauses are in fact fairly loosely related to counterpart words and thoughts. As things stand, therefore, Sellars' conception of propositions (at least as I have described it) is not appropriate for natural languages as they actually exist; it is appropriate only for a more restricted idiom that requires the functional parts of corresponding tokens to play the same semantic roles.[65] "Snow is white" and "La neige est blanche" are perhaps sufficiently similar to be semantically classifiable in the same way, as *snow is white*s, but "I will meet you tomorrow at the Blue Moon Café" and "he would meet me today at that shabby café on 14th Street" are not similar enough to be classifiable in an analogous way, so that they count as tokens of the same ostensible type. This lack of similarity does not prevent an occurrence of the latter sentence from being used to report the "content" (or be considered a semantic counterpart) of an appropriately produced occurrence of the former. Semantic counterparts are, in general, related to one another by determinate rules, but they are not always or even usually tokens of a common functional type.

What I have said about *oration obliqua* clauses also has significant implications for a current controversy about "individualism" and the meaning of a person's words.[66] According to philosophers such as Roderick Chisholm, our words have the meaning they do because of what we mean by them: our referential intentions call the tune.[67] Tyler Burge officially takes an opposite view; he argues that our words inherit their meaning from the linguistic community to which we belong. My own view is somewhere in the middle.

As I said near the end of the last chapter, I believe we gain the capacity to think (and therefore to speak) about occurrences in ancient history, exotic entities in subatomic physics, and even certain traits and quirks of everyday acquaintances only

---

[64] Sellars (1979), pp. 95-100.

[65] Jeffrey Sicha has worked out an amendment to Sellars theory that accommodates the looser relations between counterpart tokens that I have been describing here. According to his amendment, the counterpart formulas belonging to the distribution class for a propositional assertion may contain expressions with the same extensions but different meanings or implications. As far as I can tell, Sicha's amendment, which was worked out in a letter to me and is not published, successfully eludes the objections I have raised against Sellars's original theory.

[66] The word "individualism" comes from Burge (1979), whose criticism of the doctrine it applies to I shall discuss shortly.

[67] Chisholm (1997), p. 359.

by means of words that we learn from parents and teachers, textbooks and dictionaries, newspapers and television. But we do not master perfectly the words we accumulate, and our sources are also imperfect transmitters of collective verbal wisdom, so there is generally some lack of fit between our speech and thought and the speech and thought of others.

Burge supports his stronger view by arguments based on the content clauses of sentences used to ascribe those states or events to particular persons. One of his central arguments concerns a man who has an erroneous understanding of what is meant by the term "arthritis." He knows that the term refers to a painful, inflammatory disease, but he does not know that the disease is restricted to joints. He thinks he has had the disease for a long time in his wrists and fingers, but he begins to believe he now has it in his thigh, and he tells his doctor about it. It is natural to say that when he makes this statement to his doctor, he *believes* that he has arthritis in his thigh. Yet if arthritis is properly a joint disease, he misuses he word in attributing that disease to his thigh. And if he misuses it, is it not the wrong word for an accurate expression of the belief he has about his thigh?

Burge says no. In his view the man's belief is identified by means of its content, and the standard practice for attributing beliefs is to take the words a person would use to express his beliefs as *prima facie* evidence for what their content is. I think there is no doubt that if Tom candidly asserted "I have arthritis in my thigh," we would normally conclude that he said he had arthritis in his thigh and that he believed he had arthritis there. What is somewhat doubtful is the rational basis for this conclusion and the philosophical significance the conclusion should possess.

The matter of philosophical significance deserves to be considered first. Burge takes it to be considerable. He supports this verdict by describing a counterfactual situation in which the man who says he has arthritis in his thigh belongs to a community in which the word (the sound or sign design) "arthritis" has a different meaning: there it applies to rheumatoid ailments that may exist in bones as well as joints. Apart from this difference in the man's linguistic community, he is fundamentally the same; his "entire physical and non-intentional mental histories, considered in isolation from their social context," are in no way different (p. 79). But since the word he uses to describe his disease in the two cases has a different referent—an inflammatory disease of the joints in the first case and a disease of the bones or joints in the second case--his beliefs about his disease in those cases differ as well, Burge says. He will have different beliefs because his beliefs will have different contents: in the first case they pertain to one disease, in the second case they pertain to another.

To decide whether Burge is right in this last matter, we must obviously consider the justification we have for concluding that a person who utters certain words has a belief with a particular content. We certainly do not always suppose that someone who says that P believes that P. If Mary utters the words "I like Tom's sense of humor," we will unquestionably agree that she said she likes Tom's sense of humor, but we know she could be speaking ironically and possibly believe that his sense of humor is absurd or even revolting. To be convinced that she said that P we need only be convinced that she was engaging in what J. L. Austin called a "locutionary act"[68] and that she uttered words appropriately related to the words we include in the *oratio obliqua* clause following "She said that…." How she meant these words, or what thoughts she intended to convey by means of them, is not represented in our that-clause. The transition from "She said that P" to "She believes that P" properly requires information (enough to support a reasonable conjecture) about what she meant in so speaking and what her intentions were.

---

[68] Austin (1962), pp. 99-109.

Burge is well aware of this last fact, of course. To support his contention about the dependence of propositional attitudes on community standards, Burge emphasizes that partial understanding is "common and even normal in the case of a large number of expressions in our vocabularies" (p. 83) and that partially understood and even misapplied words may yet be used without qualification in giving the content of our statements and beliefs. In many of the examples he considers the subjects are willing, Burge says, to submit their statements and beliefs to the arbitration an authority, and this suggests that they are willing to have their words taken in the normal way (p. 101). This willingness, where it exists, shows the kind of intention that can often justify a transition from "She said that P" to "She believes that P." Where it does not exist, the transition is very dubiously made.

Although the acceptability of Burge's views about believing is not pertinent to my purposes in this book, I might nevertheless observe that in cases where we are intending to use a word as certain others use it, the differing beliefs that we may express with the same words in different social contexts are like the differing beliefs we may express by the words "He or she is on time" when we hear a package being left on our porch. We may have no idea who is delivering the package this time, but we use the same words anyway. In one sense we could say exactly the same thing on the following day even though, if the delivery person is not the same on both days, we would in fact be referring to different people on each occasion. Our beliefs would be semantically different each time, since they applied to different delivery-people; but they would be exactly the same in their intrinsic features. In this respect the changes of belief that Burge discusses are what have been called "Cambridge changes," changes that can be attributed to a thing because of an alteration in something else.[69] A man becomes a father when his wife gives birth to their biological child. The change he thereby undergoes is a Cambridge change, not a material change in his person.

As Burge rightly observes, "One need only thumb through a dictionary for an hour or so to develop a sense of the extent to which one's beliefs are infected by incomplete understanding" (p. 79). In spite of this incompleteness, normal speakers intend to refer to the commonly accepted referents of the names they use, and they intend to attribute to those referents, by means of predicates and associated words, the features they take to be attributed to such referents by others using those words. Who are these other persons? The variety is wider, I believe, than Burge imagines. In the first instance, they are people with whom the speaker commonly associates, such as friends or family, or they are specialists of one kind or another who may understand the words much better than the speaker does. But the speaker may also intend to conform to the usage of many special groups. Teenagers have their own *argot*, and so do philosophers, plumbers, investment bankers, and sailors. Common words are used differently in different parts of one country; and every traveler knows that many words of American English have implications that conflict in various ways with their homonyms in England, Ireland, Australia, or India. People with different education—not just in quantity but in variety—attach different meanings to different words; and the syntax of their sentences is apt to differ as well. As a result of these varying subgroups, dictionaries of different sizes give conflicting information; in some, words like "sarcastic" and "sardonic" are listed as synonyms; in others, such words are carefully distinguished.

Because different people do, in fact, mean different things by many of the lexically identical words they use, and because their referential intentions are always relevant to the interpretation of their words, Burge's anti-"individualist" arguments apply to a much smaller class of propositional attitudes than he officially supposes.

---

[69] See Geach (1969), pp. 71f.

In fact, his arguments apparently fail to rule out the individualist view of R. M. Chisholm—that the referents of a speaker's words are invariably determined by that speaker's intentions, that the words mean what the speaker intends them to mean. Since I have already expressed my disagreement with Chisholm's view, I want to say something about it.

The first point to make is that if Burge is right, Chisholm's individualist view must be compatible with the idea that speakers may misunderstand or not fully understand the meaning they intend their words to have. They intend certain words to be understood in some accepted way, but they do not fully understand what this accepted usage amounts to. Thus, the details of their meaning may elude their understanding; it is determined more by the ideas of informed speakers than by their own ideas. Their own referential intentions are important, but they do not specify the referential details.

The second point is far more important from my point of view. It is that referential intentions can give words meaning only by means of intentional behavior, and the relevant behavior is what actually gives the meaning whatever determinacy it may possess. This is true even for the incompletely understood meaning that one gives to a word one intends to be understood as it is used by certain experts. To use a word with this meaning is to defer to those experts for details about the things, events, or conditions to which they apply it. Consistency involving the details of use is what is fundamentally important—not a mere mental state of intending. Suppose a Humpty Dumpty character insisted that he can mean *horse* by "ashtray." To succeed in such a thing he would have to use "ashtray" the way others use the word "horse" or some synonym in another language. This will involve both observations and inferences. If you point to some horses in a field and ask what he thinks they are, he should say "ashtrays" if he can understand your English. If you ask him whether ashtrays can be saddled and ridden, he should assent. And so on. One can give meaning to one's words only by using them in some reasonably consistent way.

Concluding Remarks

I began this chapter by distinguishing three generic conceptions of properties, the two most important being elaborated by A-theories and F-theories, the first Aristotelian in spirit and the other Fregean or, in view of its ultimate provenance, Platonic.[70] The Fregean conception is far preferable to the other, I argued, but it needs elaboration and reconstruction if it is to be fully acceptable. Frege called his properties "concepts" (or "*Begriffe*") and, like his followers today, he thought their basic semantic function lies in relating the objects falling under them to appropriate predicates. (The predicates apply to the objects falling under their "senses.")[71] But he offered no account of this "falling under" relation and had no explanation of why object x falls under concept C1 rather than concept C2. The need for such an explanation is the basis for a long-standing problem in metaphysics, the earliest form of which arose from the assumed *chôrismos***,** or separation, of Plato's forms from the particulars we see.[72]

---

[70] After writing this I noticed a footnote in Adam's edition of Plato's *Republic* in which Adam observes that German translators of Plato's Greek mostly use "Begriff" for Plato's "*eidos*." In calling his properties "Begriffe" Frege was therefore following a familiar precedent. See Adam (1902), note to *Rep.*V, 476A2.

[71] See footnote 28 above.

[72] The problem was originally raised in Plato's *Parmenides*, 130a-135d.

I argued that predicates as well as names apply to certain objects and not others because of the way they are used by speakers and writers. Concepts as I understand them play no essential role in securing this reference. Their task is to facilitate the identification and description of predicates whose semantic features are established by speakers and writers. Like the lion, the lamb, or the llama, they are distributive objects whose ostensible traits are really traits of the objects they collectively identify or functionally describe. Although we say, "The lion is carnivorous," the predicate of our assertion is not truly applicable to an abstract object; it meaningfully applies only to living things. Something comparable is true of our assertions about concepts. We say that the concept *lion* applies to ferocious felines, but our predicate in this sentence, "applies to ferocious felines" is a string of words that, because of the way its constituents are used and structured here, strictly applies to other predicates—specifically, the word "lion," its translations in other languages, and its mental counterparts. The subject of concepts, or F-properties, is vitally important to epistemology because it shows us that a priori truth can be based on the structure of sentences (strings of words) or corresponding thoughts rather than on perceived connections between properties themselves, as rationalists perennially contend.

Chapter 5

OBSERVATIONAL KNOWLEDGE

According to classical empiricist doctrine, well-founded beliefs about matters of fact and existence can receive their support only from observation, memory, or inferences whose premises were ultimately supplied by observation or memory. On the face of it, observation is more basic than memory, because anyone who remembers that P must formerly have known that P, and when P is a contingent matter of fact, this prior knowledge could ultimately have arisen only from some kind of kind of observation. Memory and the sort of inference appropriate to matters of fact— Hume called it "experimental inference"—have always raised problems for empiricists, but observation is problematic in its own way. This chapter will deal with problems connected with the nature and scope of observational knowledge; the chapter to follow will be focused on memory, experimental inference, and the resolution of a skeptical problem that is raised in the present chapter.

A Problem about Observation

Although empiricists have always insisted that observation is our ultimate source of evidence about matters of fact, observation is actually far more problematic than it appears. In fact, it does not accord with some well-considered remarks by Hume, the most important figure in the history of empiricism. When Hume discussed philosophical skepticism in the last section of his *Enquiry*, he emphasized that the observational process typically results in beliefs or opinions, which may or may not amount to knowledge. His occasion for emphasizing this was his recollection of certain "trite topics" that skeptics "in all ages" dwell upon—specifically:

> the imperfection and fallaciousness of our organs on numberless occasions; the crooked appearance of an oar in water; the various aspects of objects according to their different distances; the double images which arise from pressing one eye... and many other appearances of a like nature.

Trite as these topics may be, they do prove, Hume admits, that "the senses alone are not implicitly to be depended on" and that we must "correct their evidence" by reason and by considerations derived from "the nature of the medium, the distance of the object, and the disposition of the organ...."[1] The evidence supplied by the senses in the observational process amounts to a thought or opinion (something propositional) because it can be corrected or corroborated by various considerations.

In these astute remarks Hume unwittingly raised a serious problem for empiricism, one that he made no effort to solve himself. The problem concerns the considerations that should be used in correcting or corroborating the thoughts generated by sensory experiences. What is the basis for these considerations? How can they be rationally supported? We should expect Hume to reply that they are supported "by experience," but it is not clear how, on his view, experience could accomplish such a thing. If the thoughts excited by a sensory experience must invariably be corrected or corroborated by empirical considerations that owe their epistemic

[1] Hume (1777), p. 151.

authority to experience, they too must ultimately have been excited by experience and appropriately corroborated. But what could have corroborated them? We seem to be faced with an infinite regress of empirical considerations or by some kind of corroborative circularity among them. The alternative of an intrinsically acceptable empirical consideration seems to be out of the question for a good empiricist like Hume.

Locke attempted to cope with the problem arising here by speaking of "the grounds of probability" bearing upon a belief or opinion. His conception of probability was primitive by modern standards, but it arose from his earnest attempt to improve upon the certainty requirement for rational opinion, which is not really appropriate, he believed, for empirical matters. Rational certainty is achieved, he thought, by demonstration or some kind of immediate insight, neither of which is applicable to uncertain matters of fact. Unlike rational certainty, which is an all or none affair, probability comes in degrees, Locke said, and it applies to beliefs or opinions created by arguments whose effect is naturally similar to that of demonstrations. In what way similar? The answer is "Similar in producing conviction." The conviction produced by empirical arguments is weaker than that produced by proofs, but it is equally concerned with truth and falsity.

The empirical arguments Locke was referring to concerned causes and effects, and the evidence appropriate to them was broadly observational. Hume would later call such arguments "experimental," but Locke did not have any general name for them. When we use them to assess human testimony about observed matters of fact, our evidence should include a variety of factors, Locke said.[2] Among them are the number of people who claim to have observed the phenomenon, their integrity as observers, and their skill in making such observations. If the testimony is taken from a book, we should consider the design or purpose of the author. As for testimony generally, we should ascertain "the constancy of the parts and the circumstances of the relation"--by which Locke probably means such things as the consistency of the testimony, its coherence, its "constancy" as to voice, competence, information, and so forth. In addition, we should consider the presence of circumstances that may have a bearing on all the preceding, such as fear, sorrow, love or hate, and the presence, nature, and circumstances of "contrary testimonies."

These grounds properly include the empirical considerations that Hume identified, because they are certainly pertinent to the truth of a claim like "I saw an oar bend when it entered the water." Locke's contention was that, to be rational, we should examine all the grounds of probability bearing on a proposition and then, upon a "due balancing of the whole, reject or receive it with a more or less firm assent proportional to the preponderancy of the greater grounds of probability on one side or the other."[3] In addition to wondering how a ground of probability can be also wonder about the principles to be followed in balancing probabilities. Neither Hume nor Locke identified these principles, and it is hard to see how they might have proceeded to do so. However that might be, we have the surprising result that, according to fairly explicit testimony that can be found in the writings of both Locke and Hume, observational evidence should be assessed by general considerations of an empirical kind. This is surprising, because classical empiricism is generally associated with the doctrine that general beliefs about the world must be founded on a generalization from experience. This testimony seems to turn things upside down.

---

[2] Hume discusses the way his thinks human testimony should be critically assessed in the second part of his chapter on miracles; his views are generally similar to Locke's but they are directed to a particular topic, miracles, and they are set forth less systematically. See his *Enquiry* of Human Understanding, Section X, Part II.

[3] Locke (1984), vol. 2, Bk. IV, ch. 15, sect. 5, pp. 366.

Evaluating Observational Beliefs

If we think about the considerations Hume and Locke mentioned in relation to the assessment of human testimony or observational evidence, we can identify four basic sorts of things that should be weighed when we are evaluating a belief to the effect that some person perceives something. Suppose, for example, that our friend Tom forms the belief that he sees a gray cat in a nearby room. If we are seriously concerned to ascertain whether he does indeed see such a thing,[4] we should begin by reflecting on the perceptual process he employed--in this case, seeing. Tom may or may not be good at this process. His vision may or may not be acute; he may or may not be able to distinguish grays from tans; and it is even possible that he is blind and responding, in the present case, to a hypnotic suggestion. Another matter is the nature of what he ostensibly observes. Are there such things as gray cats? Are cats easy to see? Are they large, small, opaque, or transparent? Are there other things that look like them--things that a person like Tom might naturally (in these circumstances) take to be a cat? Then there is the character of the circumstances (the conditions) in which Tom's belief is formed. Was the light good enough for a person, a person like Tom, to see a cat, a gray one, at the distance in question? And finally there is the sort of observer (the sort of person) Tom is. Is he intelligent, sane, critical--or gullible and demented? Is he obsessed with cats? And so on and so forth.

Some of these questions might be answered by observing Tom's behavior and recalling how he has behaved on various occasions in the past, but these answers will raise further questions of the same kind about our own observations. Even if we could answer these further questions without circularity, we would be faced with general questions about observers, observable objects, observational processes, and conditions of observation whose answers are relevant to the evaluation of *anyone's* observational beliefs or reports, ours or the wisest and most perceptive among us. But to support these answers by reference to anyone's observations would be reasoning in a circle.

Empiricists rejecting Hume's "mitigated" skepticism often felt justified in avoiding these questions because they were convinced that empirical knowledge rested on a special kind of observation that rendered them unproblematic. This kind of observation might better be called "direct apprehension," since the objects it is concerned with are supposed to be subjective sensory objects. In chapter one I mentioned that Bertrand Russell considered our basic empirical knowledge to be obtained by "acquaintance," a process he understood to involve a direct confrontation with the objects of our immediate experience. He called these objects "sense data,"[5] but they could equally be described as *sensa* or what we sense when we have sensory experiences. Obvious examples of sense data are itches, feelings of pain, afterimages, and the supposed sensory objects involved in the experiences we have when we look at purple clouds, fall foliage, or indeed when we have any perceptual experience. Since these supposed objects were thought to be directly and wholly presented to a subject, they were considered to be what the subject perceived them to be. No empirical assumptions relating to their nature, the nature of the process by which they are apprehended, the nature of the circumstances under which they happen to be apprehended, and the nature of the apprehending subject (the person) were considered pertinent to what the subject knows in apprehending them.

---

[4] The question is whether he sees such a thing, not whether such a thing exists. The answer to the latter might be used as evidence for the answer to the former, but the questions are nevertheless different and should be answered in different ways.

[5] Russell (1953), p. 198. For a very perceptive up-to-date discussion of sense data, see Huemer (2007).

I have spoken of sense data as supposed objects because Russell's successors soon came to doubt that such things actually existed.[6] The doubters generally conceded that we have all sorts of sensory experiences, but they argued that sensory experiences were states of sensing that did not include the sensed objects that Russell called sense data. The arguments they offered for this surprising negative view—surprising, because it seems obvious that we do apprehend something subjective when we have after-images or double vision—were based on general considerations of a theoretical sort. Perhaps the decisive negative argument, the one most effective in convincing philosophers to reject sense datum theories, was to the effect that accepting such theories is tantamount to asserting that a sensuous curtain stands between perceiving minds and the world they normally believe they are perceiving.[7] This sensuous curtain shields the external world from our perceptual activity and renders it fundamentally unknowable. It becomes an incomprehensible Kantian "thing-in-itself."

Philosophers rejecting sense data on these grounds often thought they did not have to cope with the questions I asked about perception because they were convinced that sensory experiences bereft of sense-data could yield knowledge without presupposing empirical information about perceivers, perceived objects, the process of perceiving, and the effect of background conditions on what subjects might suppose they are perceiving. Roderick Chisholm stubbornly insisted on this, and he convinced many philosophers that he was right. In perceptual experience we are "appeared to" in certain ways, he said, and being appeared to is a "self-presenting state," one that necessitates the certainty, for the subject, that he or she is in that state.[8] Since being certain about the state one is in involves certainty about what that state is like, one's certainty about the character of one's current sensory experience depends only on that experience, not even partly on the empirical assumptions I have described.

Chisholm's position on this matter is far from convincing. Although some sensory experiences do, at least in some circumstances, seem to present themselves to our consciousness,[9] there is no good reason to suppose that we *cannot* make errors about them. Our access to them may be privileged, but it is hardly infallible. There is good evidence for this. Scientific studies have shown that people make all sorts of errors about the character of their mental states and sensory experiences; and they are often entirely unaware of what they are thinking, sensing, or feeling, particularly when their attention is focused on something else. The errors they make about their sensory experiences do not depend on a particular conception of them—whether they are understood as involving sensory objects (sense data) or not. People are simply not infallible in identifying, describing, or otherwise ascertaining the specific character of their sensory experiences or their conscious states generally.[10]

---

[6] See Barnes (1944-45), Quinton (1955), and, again, Huemer (2007).

[7] Ayer (1956), p. 117.

[8] See Chisholm (1976), p. 26.

[9] Normal human beings have sensory experiences whenever they perceive anything, but they are rarely aware of their sensory experiences, having their attention focused on things and persons in the world around them. For discussion, see e.g. Johnson (2006), BonJour (2007), Crane (2006).

[10] I discuss some pertinent studies in Aune (1967), pp. 31-38; other studies concerned with reports on mental processes generally are discussed in Nisbett and Wilson (1977), a study every philosopher should be familiar with. See also Williamson (2000), ch. 4.

Does Knowledge Need a Foundation?

The idea that we could make errors about the character of our sensory experiences—as opposed to the physical realities whose existence those experiences normally indicate—might seem perplexing or even alarming to philosophers who assume that if we have any empirical knowledge at all, it must rest on a foundation of something directly knowable. Classical empiricists commonly made this assumption. The basis for it is a regress argument that goes back to Aristotle.[11] According to this argument, if an empirical fact is known by means of some inference, the premises used in the inference must be known to be true. If those premises cannot be known to be true non-inferentially—by some kind of direct inspection—they will have to be known by reference to some more basic premises, which will have to be known to be true as well. Since one cannot possibly know something P on the basis of knowing some Q that is knowable, ultimately, on the basis of P itself, either the regress stops with *some* non-inferentially known fact or facts or P is not really known at all. Since nothing is a better candidate for being directly knowable for a subject S than the character of S's own experience, knowledge of such experience must be knowable in that way.

This argument is irresistible if we suppose that we have empirical knowledge and also believe that the certainty requirement is applicable to it. According to that requirement, anything we actually know is either immediately certain (certain without reference to anything else) or a provable consequence of other things that are immediately certain. But as I argued in chapter one, this requirement is not applicable to routine examples of empirical knowledge. According to the standards we normally use in everyday life, we know many things that are not immediately certain or provable consequences of other immediate certainties. I know that I live in the State of Massachusetts, but my knowledge of this, well-supported as it is, [12] is not an immediate certainty or something I have inferred from immediate certainties. Thus, when existing standards are assumed, the regress argument has no force and there is no plausible basis for inferring that our empirical knowledge, all of it, rests on some directly knowable foundation of certain truth.

If we allow that empirical knowledge nevertheless requires some starting place, one that may be neither indubitable by the subject nor the result of some actual inference, we still do not have to agree that it consists wholly or even partly of facts about the knower's sensory experience. What would suffice is some report or assertion (verbal or mental) that is reliably correlated with the sort of occurrence or state that makes it true. The man's assertion that the bird whose cry he hears on a remote lake in Maine is a loon has this kind of reliability, and so does the woman's assertion that the façade she sees through the window of a train (when no barn-façades are in the area) is that of a barn. In both cases the observers would normally be taken to know what they say they are hearing or seeing. Their evidence in these cases is, of course, defeasible and could therefore be overridden if countervailing evidence should become available—evidence about the prevalence of phony barn façades and evidence about fake loon calls broadcast on northern lakes by scientific maniacs with powerful amplifiers. Yet in the absence of such evidence there would normally be no question that the observers have the relevant knowledge.[13] This evi-

---

[11] Aristotle, *Posterior Analytics, 72b18.*

[12] I have, of course, an enormous amount of evidence for it—so much so that I can hardly survey it all or identify the most important elements belonging to it..

[13] Huemer (2001) says that observers' "seemings" that they are perceiving a tree or house are "presumed true, until proven false' (p. 100). BonJour (2004) rejects the direct realism of Huemer and other philosophers, claiming that they "offer little or nothing by way of a positive account of how perceptual beliefs are

dence falls short of what is required by the certainty requirement for knowledge, but when that requirement is applied there is almost no room for empirical knowledge at all.[14]

It is important to realize that the existence of stopping places or non-inferred items of ostensible knowledge does not imply that persons having it need not possess knowledge of other, related things. If the man in Maine did not know what a loon is, did not know they cried in a striking way as they flew about the lakes they inhabit, he would not know that he was hearing a loon; and the woman would not know that the façade she sees is that of a barn if she did not know what a barn is and what it is typically used for. But this background knowledge need not provide premises from which the reports about the loon and the barn were inferred. To have imperfect knowledge—the kind not requiring rational certainty—a subject typically has to satisfy a cluster of epistemic conditions whose general character I discussed in chapter one, but these conditions do not require that some inference be made. Some of the required background knowledge may be more general, moreover, than the knowledge provided by observational reports. The structure of empirical knowledge is therefore quite different from the one suggested by Aristotle's regress argument.

The fact that ascriptions of knowledge are commonly assessed and even made on the basis of defeasible presumptions about the causes and effects of familiar phenomena suggests a solution to Hume's problem about how the evidence of our senses could possibly be corrected by reason and empirical facts about perceivers and perception. For any empiricist, reason itself—or pure reason, as Kant described it—has a very limited role in assessing observational claims. If these claims are to be assessed by background information, that information must have arisen empirically from observations of a less critical, or less cautious, kind. This is no doubt the way we got the information in the first place. We began innocently and naively, but we soon became more critical. Conflicting testimony required us to change our minds in many instances, and our naïve presumptions gave way to corrected and improved ones.[15]

Here is a simplified picture of how the correction and improvement is accomplished. There is an initial presumption, tacit rather than explicit, that able-bodied people who are reasonably young can equally discern what is present to their senses. This presumption is not baseless; it is supported by the agreement such people commonly reach about what visible, audible, fragrant, or foul. But this agreement is imperfect; the parties sometimes disagree about what is discernible under these or those conditions. Fortunately, regularities occur among the occasional disagreements, and we eventually conclude that, just as some people are stronger or can run faster than others, some people are better than others at seeing, hearing, smelling, or tasting. We also conclude or, better, learn that some perceptual conditions facilitate while others hinder the identification of colors, sounds, tastes, and smells. As we reflect on conclusions of this kind, we draw distinctions and adopt explanatory hypotheses that correct our original presumptions. We now consider some observers more reliable that others; we now regard some perceptual conditions as optimal for certain sensible qualities; and so on.

The generating conditions for the higher-order knowledge needed for the assessment of ground-level observations is thus human inconsistency and judgmental

---

[actually] justified according to their view." I agree that more can and should be said about this matter; I make some remarks about it later in the chapter and near the end of chapter six.

[14] See Fogelin (1994), p. 140.

[15] The idea that what might be considered initial opinions give way to improved ones in the process of rational inquiry is nicely modeled by Gupta's discussion of interdependent definitions in Gupta (2006), chapter 3.

conflict. Whether we are concerned with a single person's experiences or with intersubjectively available external objects, inconsistent reports are not only possible but actual. We resolve the inconsistencies by drawing distinctions and adopting explanatory principles, which we may eventually have to revise again. Sometimes our predications fail, and other explanations come to mind. We nevertheless become accustomed to resolving conflicts by rejecting some observational reports in favor of others. We disallow some because of the conditions in which they are made; we reject others on account of an ostensible defect in the perceiver's sense organ; and we disallow many because of the perceiver's distraction, lack of concentration, or carelessness and inattention. As our knowledge grows in respect to observers, ways of observing, observable objects, and conditions pertinent to observational success, we naturally take account of an ever-wider range of phenomena when we assess observational reports. Background theory thus becomes increasingly important. Observation loses its autonomy and becomes subject to higher-order principles.

The new principles we introduce in coping with observational conflicts do not concern only outer things; they also concern the nature of our experience when we make observations. The red, green, or gray things we perceive are located in space some distance from us; but sometimes we perceive something we want to describe with these color-words when nothing so describable available externally. These anomalies prompt us to think of our perceptual experiences as occurrences that somehow have qualities themselves. Wilfrid Sellars famously used a myth, the myth of Jones, to explain the origin of human discourse about sense-impressions,[16] but the concept of a sense-impression is not something every philosopher, let along every intelligent adult, will acknowledge having. So-called disjunctivists about the objects of experience say that we either sense external things or suffer hallucinations, but we do not sense something inner or subjective (objects or processes) in both cases: when we perceive something we do not sense something in addition to what we perceive.[17] The metaphysics of experience is now a contested subject,[18] and I lack the space to pursue it here. But I have no doubt that we do have perceptual experiences when we sense external objects and that these experiences possess qualitative features of their own—features that we can normally describe only by words strictly applicable to external things. One morning in the distant past I had the sort of experience spiritualists describe as seeing an apparition. I have always thought of it as a hallucination, the only hallucination I can remember ever having. My only way of describing it is this: "It was an experience of ostensibly seeing a drab motionless woman suspended in front of my bedroom window."

There is no primitive stage of our intellectual development when we did not think of ourselves as embodied creatures perceiving one another amid the objects of a common world. Our talk of our "selves" makes clearest sense only in connection with the thinking animals we actually are. But we normally look away from our selves when we perceive things, and not perceiving our arms, legs, or noses, we can be tempted to think of ourselves as something very different from an embodied creature. We might even come to believe, as Roderick Chisholm did, that we are spirits who move about the world and perceive it only "by means of" the body we inhabit and such things as the glasses that are perched on the body's nose.[19] But our ultimate evidence for our beliefs about ourselves and our world is the experience we

---

[16] Sellars (1959).

[17] See Crane (2006), section 3.4, and also Johnson (2006), pp. 286-89. A sense-impression as Sellars describes it may not be an object of inner sense, but disjunctivists would almost certainly reject such a thing anyway, as Johnson (2006) appears to do on p. 288.

[18] See the essays in Gendler and Hawthorne (2006).

[19] Chisholm (1991), p. 171.

have, not our inner states. That experience, which is fundamentally focused on the "objects" of our senses, needs to be understood theoretically, like anything else. Its character and place in the scheme of things needs to be investigated. As we pursue this investigation, our conception of our experience and our selves develops and changes. Current debates in the philosophy of mind show that this development is far from over. What we know about our sensory experience is therefore modified by higher-level inferences; it does not provide an immutable foundation on which the rest of our knowledge rests.

Alternatives to Foundationalism

The structure that I have just described does not accord with the usual alternative to the view, call it "foundationalism," that knowledge rests on a foundation of uninferred certainties or, as some say,[20] likely truths. The usual alternative is coherentism. According to this view, only beliefs can add credibility to beliefs,[21] and they can do so only in the context of a larger system of beliefs, one in which each member gains an indirect justification from the size and coherence of the system as a whole. The notion of coherence involved here is explained in different ways by different coherentists, but the alternative explanations are generally similar. BonJour, who was once a coherentist,[22] said that a coherent system must be consistent, both logically and probabilistically, and that its consistency is enhanced by inferential connections between its constituents and diminished both by unexplained anomalies among them and by relatively autonomous subsystems including them. [23]

The interplay between the justification a belief receives from another belief and the justification both beliefs indirectly receive from the system to which they belong obviously requires very careful treatment. How large must a system be to make a given belief strongly justified, all things considered? How large and how coherent must it be if a given belief amounts to knowledge? Fogelin (1994) once asked rhetorically if any human system of beliefs has ever satisfied BonJour's standards for coherence, and he answered in a way suggesting he thought the answer is clearly no. If this answer is right, as I think it is, presumably no belief has ever been justified and no one, by BonJour's coherentist standards, has ever known anything. This would have been an exceedingly unwelcome outcome for BonJour when he was a coherentist, because he developed those standards as a means of avoiding skepticism[24].

Although BonJour was convinced that foundationalism had unacceptable consequences, his coherentism involved the "internalist" conception of epistemic justification that was central to foundationalism. According to this conception, if belief A is not self-justified, it is justified by an inference from some belief B.[25] I have not explicitly criticized this conception of justification; I have simply not used it. In speaking of knowledge I have instead followed Lewis and spoken of evidence.[26] I could just as well have spoken of justification, but if I had done so, I would not have un-

---

[20] BonJour 1999), p. 230.

[21] See Davidson (1986), p. 311.

[22] He abandoned coherentism in BonJour (1999), where he defended a form of foundationalism. An unrepentant coherentist, whose views are far more complicated than BonJour's, is Lehrer (1997). My reason for rejecting any form of coherentism appears at the end of this section. But also see BonJour (1999).

[23] BonJour (1985)

[24] BonJour (1985), p. 80.

[25] Not all philosophers who are internalists about justification would accept this. Feldman speaks of a feeling of wamth as a reason for believing something, but feelings are not premises from which conclusions can be inferred. See Feldman (2005), p. 273.

[26] See chapter one, pp. 13f. The notion of justification did tacitly come into my account of knowing for certain. See my definition on p. 34 and the paragraph immediately following it.

derstood it as most internalists do.[27] The man who identified the loon's call made no inference; his belief that he was hearing a loon could be described as justified in the circumstances but not justified by itself: the available sounds are vitally important. In rejecting the typical[28] foundationalist's assumption that a belief is either self-justified or justified by means of another belief, I am therefore also rejecting the coherentist's assumption that a belief can be directly justified only by another belief.

In spite of this dissent from the internalist assumptions of the typical foundationalist and the coherentist, the picture of empirical knowledge that I favor might be described as a picture, really, of organized belief. A philosopher committed to what I called the certainty conception of knowledge might in fact insist on this description. That is all right by me. We can of course—within limits—use the word "knowledge" as we want, and if we want to restrict knowledge to instances of rational certainty, we can certainly do so. But as I argued in chapter one, existing usage does not impose this certainty restriction on the word "knowledge." There is therefore nothing incorrect in speaking of knowledge as I am doing here. I shall later, as I implied at the end of chapter one, discuss reasons for occasionally seeking greater certainty than what we ordinarily regard as satisfactory, and I shall treat these reasons sympathetically. But inherently uncertain "knowledge" is what we ordinarily have, seek, and argue about. The structure of that knowledge is what I have been discussing.

It is the fallibility of the best judgments we usually call knowledge that destroys any supposed invariant foundation for empirical knowledge. As we gain information about the nature of perceivers, perceived objects, perceptual processes, and background conditions that affect perception, our assessment of particular observation claims becomes so theory-dependent that we cannot realistically isolate an independent "observation language." In fact, in continuing to learn details about the microstructure of our world, we can easily reach the point of using paradigmatically theoretical language in making routine observations. This language may be highly exotic in university laboratories, but even wags in high-school lunchrooms have long been wont to speak of drinking H_2O or seasoning a dish with a little more NaCl.

These last observations might seem to support coherentism, but they do not actually do so. Although observational knowledge increasingly involves higher-level principles, the totality of what we know, or think we know, typically contains a lot of disorder.[29] This disorder is increased if we think of the knowledge of a single person rather than (as often in philosophy) "our" knowledge, the knowledge of some idealized community. Even in the best universities, physicists may be badly informed about the latest developments in psychology or molecular biology; and mathematicians or philosophers may know next to nothing about diplomatic history or agronomy. A plain person, one who is not a professional knower, may have a real hodgepodge of knowledge; the aggregate will depend heavily on special interests, such as photography or the propagation of Hosta lilies, and on how much reading the person has done, and in what subjects. When we ask the impersonal question "What is known about the structure of space or the interface between chemistry and physics?" we may learn that what the best and brightest collectively know about these matters is very well organized, but there are gaps in even collective wisdom, and knowledge in some areas—for instance, the breeding of tigers in captivity—may have very little to do with any scientific discipline. Far from being an organized system growing from some single source, knowledge in the sense of what is known empirically has

---

[27] See footnote 25 above.

[28] BonJour is now a foundationalist who holds that a basic belief can be justified by a perceptual experience, which is not and does not include a belief. See BonJour (1999), p. 230

[29] Gupta (2006) describes our view of the world as "a collage of conflicting pictures" that "contains both empirical anomalies and conceptual paradoxes" (p. 200).

no general, specifiable structure. It is something of an aggregate of aggregates with a jungle of twisted and gappy connections.

These last observations apply to Quine's "holism" as well as to coherentism. In "Two Dogmas…" Quine famously said:

> The totality of our so-called knowledge or beliefs, from the most casual matters of geography and history to the profoundest laws of atomic physics or even of pure mathematics and logic, is a man-made fabric which impinges on experience only along the edges…. A conflict with experience at the periphery occasions readjustments in the interior…. Any statement can be held true come what may, if we make enough adjustments elsewhere in the system.[30]

Although Quine was disputing the supposed autonomy of analytically true statements in this passage, his words have been taken, rightly or wrongly, to support the idea that our beliefs form an interconnected web, elements of which are indirectly supported by the degree to which the whole structure comports with nonverbal experience. As I shall show in the next chapter, however, empirical confirmation is really not this holistic. Individual statements are always confirmed together with some others, but the totalities thus confirmed are not as extensive as Quine's words suggest. Just think of how you might confirm Tom's belief that it is freezing outdoors. Seeing that there is snow on the ground and that the plate of water left out for the birds is now a plate of ice would normally be sufficient; it would not require any consideration of evolutionary theory or the laws of supply and demand. Sometimes whole theories may be relevant to the confirmation of some empirical statement, as it often is in subjects such as astrophysics,[31] but this is far from usual. The whole of twisted and gappy connections I described above is actually compatible with everything we know about empirical confirmation.[32]

Knowledge and World: Some Problems

What I have been saying about knowledge here applies to imperfect knowledge, or knowledge ordinarily understood. As I have observed, this knowledge is based squarely on defeasible presumptions. Although these presumptions, which are known to be generally reliable, are commonly accepted and rarely questioned in everyday life, they provide well-known targets for philosophical criticism. The evidential basis for the criticism is sometimes very reasonable, and it deserves to be met. I will discuss some of it in the rest of this chapter.

A basic theme in much of this criticism is the idea, which I have just been criticizing, that our beliefs about the world around us must ultimately be inferable, if they are rationally defensible, from the character of our sensory experience. Apart from the claims I mentioned earlier—about what is directly and primarily knowable, empirically—two additional reasons are often given for this idea. The first pertains to the transmission of information. Although it may appear that we are in direct contact with the physical things we perceive, there is actually a significant logical gap between the information we receive and the spatially separate objects transmitting it. To have actual knowledge of those objects, there must therefore be a flow of information from them to us: we must absorb that information and consciously take account of it.[33] Doing this, whether we consciously realize it or not, requires understanding and rational principles. We must in fact draw conclusions about what is ex-

---

[30] Quine (1953), p. 42.
[31] See Gribben (1998), p. 184.
[32] See chapter 6 below.
[33] See Dretske (1981), ch. 6.

ternal from information that is produced within us.[34] At the very least, we must infer causes from received effects.

The second reason frequently given is that the external objects we might think are presented to us in experience are actually very different from anything that is presented there. Descartes was the first philosopher to emphasize this point, insisting that external objects could be exhaustively described in relation to their geometrical and kinematic qualities.[35] We normally think of external objects as colored, noisy, or fragrant on account of the effects they ultimately have on our consciousness, but these effects, which Descartes regarded as ultimately determined by the geometry of external objects, light, and our sensory receptors, are ideational states productive of beliefs about those objects. Our current scientific beliefs about the external world are, of course, not the same as those of Descartes, but they agree with his in not ascribing the sensuous qualities we discern in our perceptual experience to the objects themselves. Our sense organs and nervous system contribute to their character just as much as their more remote external causes do.[36]

Although both these beliefs have historically led to skeptical doubts about the very existence of a world external to our consciousness, it is obvious that one of them is based on the idea that we know the external world exists and also know a lot about its nature. A serious epistemological question that may yet be asked is "How is this knowledge possible?" or, less cryptically, "How can we possibly know what we think we know about this so called external world?" Hume in effect raised this question and came to the conclusion that it cannot be answered affirmatively: this supposed knowledge is not possible; we cannot really have it. His reasons for this conclusion are worth recalling, because they are still relevant to philosophical thinking on the subject of the external world. Some philosophers have recently defended principles that are tantamount to the ones Hume assumed.[37]

One of the reasons Hume gave involved a particular conception of legitimate non-deductive inference. What he called "experimental inference" is causal inference, or inference relying on a causal principle such as "Scratching dry, well-made matches on a rough surface in the presence of air causes them to light." Inference relying on such principles—for instance, an argument concluding that a match satisfying the conditions mentioned in the causal principle will, having been scratched, light—would now be considered deductive rather than experimental, but the cognitive process Hume described as giving rise to our belief in these principles would be considered inductive today. As Hume explained it, this process involved a generalization from experience. On experiencing a "constant conjunction" between occurrences of a kind A and a kind B, we form the belief that B-occurrences are caused by A-occurrences, and the strength or firmness of our belief is determined, he said, by the number of these transitions that we experience. Hume did not actually describe the process of forming these beliefs as an inference because he could not specify an appropriate rule of inference. He thought we simply and naturally formed the appropriate belief when the constant conjunctions we experience become "sufficiently" numerous for us. The number of conjunctions required in the case of this or that person is purely an empirical matter.[38]

In spite of Hume's celebrated doubts about the rationality of this belief-forming process and the relative frequency of obtaining true beliefs by means of it, he considered it an acceptable process of belief-formation, the only such process

---

[34] BonJour defends this nicely in his 1999 essay.
[35] I discuss Descartes' view of the external world in Aune (1991), chapter one, section 6.
[36] See the Introduction to the essays on the science of color in Byrne and Hilbert (1997).
[37] I show this in chapter six, when I discuss problems about inductive inference.
[38] I discuss Hume's epistemology in some detail in Aune (1991), ch. 3.

available to us. But if we need to infer external causes for our subjective experiences by a process of this kind, we cannot possibly succeed. To employ the process we will have to experience a constant conjunction between our subjective experiences and their external causes, and to do this we will have to experience those causes directly—and this, Hume thought, is something we cannot do. If we could do it, we would have no need to infer their existence by any kind of reasoning.

The impotence of Hume's form of experimental inference to justify our supposed knowledge of the existence and nature of an external world does not necessarily raise a problem for contemporary empiricists, because many of them who think that the existence of external objects needs to be inferred by some kind of non-deductive reasoning accept forms of inference that do more than generalize from experience. One currently favored form, used by BonJour in justifying his belief in a world external to his consciousness, is Inference to the Best Explanation.[39] As it happens, there are serious problems with this form of inference; so one current way of providing such a justification remains questionable. (I shall discuss this matter further in the next chapter, where I discuss various forms of experimental inference.) But Hume had, as I said, other reasons for thinking that we cannot really know what we think we know about the external world, and one such reason has been given in a recent argument purporting to defend a conclusion directly contrary to Hume's. In Hume's thinking this reason supported a serious doubt about the meaningfulness of talk about a domain of objects that we cannot directly observe. The recent argument supports a similar doubt, but it is intended to undermine the kind of skepticism that Hume espoused.

It was a basic tenet of Hume's philosophy that meaningful words express genuine ideas and that genuine ideas arise from experience. Hume called the experiences from which genuine ideas arise "impressions" and claimed that any ostensible idea must, to be genuine, be derivable from one or more impressions. To be derivable from a single impression an idea must be a copy of that impression; to be derivable from a group of impressions, an idea must be complex and each idea ingredient in it must be a copy of some impression. The impressions copied by a person's genuine ideas must, of course, be impressions that person has actually had. Every genuine idea is the effect of one or more antecedent impressions. Since the supposed idea of a world external to one's consciousness could not have arisen from internal impressions, this supposed idea is bogus and cannot confer meaning on any word. If the term "external world" is meaningful, it cannot therefore have the meaning it seems to have; it cannot refer to anything that does not belong to a person's experience.

By current standards Hume's principle of meaning sounds very crude, but it was taken up and polished by nearly two hundred years of empiricist activity. The concept of experience underwent significant changes during this period, and the required connection between experience and meaning changed significantly as well. Some early twentieth-century empiricists held that empirically significant statements must be "reducible" to statements that can be verified by experience; logical positivists maintained that they must themselves be so verifiable. As I noted in the preface to this book, logical empiricists repudiated "reductionism"; they abandoned verificationism for confirmationism, the thesis that meaningful sentences must, at least in the context of some theory, be subject to empirical confirmation: they must support predictions that could in principle be verified and, if verified, would increase their probability. The last attempt by a logical empiricist to work out a satisfactory confirmation criterion of empirical meaningfulness was recorded in Rudolf Carnap's

[39] Se BonJour (1999).

"Methodological Character of Theoretical Concepts."[40] In 1958 David Kaplan discovered a problem with the criterion Carnap offered in this paper, and Carnap subsequently abandoned the project.[41] Evidently he did not think an acceptable criterion for empirical meaningfulness could be found.

Although logical empiricists accepted Carnap's verdict on attempts to specify a criterion of empirical meaningfulness, the one-time empiricist Hilary Putnam has recently defended a new principle of meaning closely related to Hume's principle. Putnam's principle is a version of what is known as "Semantic Externalism," and it has a very positive bearing on Hume's problem about the external world. Putnam introduced his principle in an effort to show that Hume's problem cannot meaningfully arise. It is arguable, however, that Putnam's principle has a general, unsatisfactory consequence that was characteristic of Hume's principle: we cannot meaningfully say what we want to say, or think we are saying, about domains to which we lack experiential access. Putnam supported his principle by a now-famous thought experiment about brains in a vat, a thought experiment that adds vitality to skeptical doubts about a world external to our consciousness. It raises a problem even for philosophers who wish to maintain that we directly perceive an external world.

Semantic Externalism

Putnam introduced his semantic externalism by commenting on a famous paper by Alan Turing. In 1950 Turing proposed what he called an "imitation game" as a means of determining whether an appropriately programmed computing machine could reasonably be considered conscious. His idea was that if a scientific investigator, having examined a sufficient number of typewritten responses to questions designed to determine whether the respondent is an intelligent human being or a computer programmed to mimic the responses of a human being, could not distinguish the human respondent from the computer, then the investigator would be entitled to conclude that the computer is a thinking thing.[42] Putnam, after describing Turing's imitation game in some detail, asked whether a similar test could determine whether the words produced by a machine would actually refer to what a person using those words would be referring to. Putnam's answer was no. However natural and well composed the machine's responses to the investigator's questions may be, if the machine has no sense organs or other hookup with the objects it is ostensibly writing about and no motor organs for interacting with those objects, it will not, he said, be *referring* to anything at all. If it is merely playing an imitation game, the words it produces will in fact be comparable to the sounds produced by a record player: they will not be genuinely referring uses of language.

Without attempting to specify the minimum conditions necessary for genuine reference, Putnam proceeded to apply his general conclusion about reference to a hypothesis that a skeptic might cite in support of a negative assessment of ordinary perception. The hypothesis concerns a number of brains immersed in a vat of nutrients and connected to a super computer in such a way that they have the sensory experiences of seeing, hearing, smelling, and physically interacting with a spatially extended external world of objects and persons. According to the hypothesis, the brains will believe they inhabit the spatio-temporal world we believe we are experiencing, but they will be wrong. The world of their experience will be a delusional

---

[40] Carnap (1956).

[41] To my knowledge, Kaplan never published his criticism, but he described Carnap's reaction to it in Kaplan (1971).

[42] Turing (1950). Actually, Turing used the question whether the investigator can distinguish the computer responses from the human responses as a replacement for the question "Can machines think?" He thought the latter question was "too meaningless to deserve discussion."

world; their reality will be utterly different from what they think it is. The same could be true of us, the skeptic says. For all we know, we ourselves could be similar brains in vats. If we cannot eliminate this possibility, we cannot reasonably contend that we are what we think we are and that we know what we think we know.

Putnam used his thesis of semantic externalism to attack this skeptical scenario. Since the brains in the vat of nutrients are not and, he assumed, never have been in causal contact with a world of objects external to their consciousness, their thought-words cannot actually refer to such objects as vats, trees, and other persons. If their thoughts refer at all, they refer to the entities that stimulate those thoughts: their subjective experiences or elements of the computer programs that produce those experiences. This fact about reference undermines the skeptic's contention, because it implies that a sentence or thought "I am a BIV (a brain in a vat)" could not possibly be true. If a subject thinking this thought could, by means of it, think about actual brains in vats—that is, mentally refer to them—the thought would be false: the subject's thoughts and experiences would be connected to an external world containing vats and other things. If, on the other hand, the subject were not connected to real external objects, the sentence or thought "I am a BIV" would not refers to vats and could not thereby say something true about them. In one way or another, therefore, an utterance or thought "I am a BIV" could never truly affirm that the subject is a BIV. In consequence, it could not support a genuinely skeptical hypothesis.

Putnam expressed the conclusion of the argument in stronger terms than I have used here; he said we can *know* that we are not brains in a vat. This stronger conclusion does not appear to follow from the premises of his argument, however. Consider the assertion "I am a BIV." According to the argument, if I am a BIV my words "I am a BIV" do not have the reference they appear to have; they refer to subjective experiences or features of a computer program. If I am not a BIV but a rational animal, then my words do refer to a BIV and are false. Since I am either a BIV or a non-BIV, it follows that my words "I am a BIV" either have an exotic meaning or they assert something false of a real person. But neither disjunct of this consequence is shown to be true by Putnam's argument, and the disjunction as a whole does not entail that I have the categorical knowledge Putnam says I have.

If I had some direct knowledge of what my words (or thoughts) refer to, I could eliminate one of the disjuncts in question, but if Putnam's thesis of Semantical Externalism is true, direct knowledge of this kind is out of the question since it depends on causal factors external to my consciousness. I cannot therefore argue:

> My words "I am a BIV" refer to an embodied me and a real external vat. Therefore my words or my thought do not have the exotic meaning mentioned in the last paragraph. Therefore, my words "I am a BIV" assert something false about a real person. Therefore a real person and a real vat exist. I know this last fact. Therefore, I know that I am not a BIV.

My epistemic predicament appears to be described by the disjunctive conclusion that either my words "I am a BIV" have an exotic meaning or they assert something false of a real person. I do not know what my words actually mean and I therefore cannot eliminate either disjunct and so ascertain my true status as a thinking being.

In an encyclopedia article Anthony Brueckner suggested that the conclusion Putnam wanted can be obtained by a variant argument:[43]

[43] Brueckner (2004).

a. If I am a BIV, then it is not the case that if my word "tree" refers, it refers to trees.
b. If my word "tree" refers, it refers to trees.
c. Therefore, I am not a BIV.

The first premise here is supposed to be a consequence of SE, Putnam's semantic externalism. The second premise is supposed to state a semantic fact that speakers can know a priori about their language, whatever it is and wherever they are speaking it. Thus, by virtue of knowing what "refers" means and knowing the meaning of quotation marks, speakers can supposedly know that disquotation is applicable to any successfully referring expression of their language. Since these two premises entail that the conclusion C is true, any speaker or thinker to whom "I" applies can supposedly know that he or she is not a BIV.

The argument is not satisfactory because a speaker (or thinker) to whom the disjunction I mentioned applies would not know what either premise is referring to. Suppose the referent of "I" is a certain BIV. The premise will then be true, but given SE the speaker could not understand what it is supposed to say—could not think the corresponding thought—because the speaker cannot comprehend a reference to trees. Similarly, if such a thinker entertained premise B, it would be thinking, "If my word 'tree' refers, it refers to trees*," the asterisk implying that the subject is thinking about what Putnam calls "trees-in-the-image," not trees in the intended sense. Non-BIVs could, of course, express the thoughts appropriate to the premises and conclusion, but if SE is true, they would not know what thoughts they would be expressing and so would not know that C is true.

The idea that we do not have a direct, privileged access to what our words or, more generally, our ideas refer to is contrary to standard empiricist doctrine, but Putnam accepted it, saying "meanings just aren't in the head."[44] If Putnam is right about this—if the meaning, the referential character, of a word or idea is in a significant way determined by input/output causal relations holding between that word or idea and objects in the world—then the empiricist idea that analytic truths do not (as Hume put it) depend on anything that is anywhere existent in the world must apparently be false. Putnam's Semantic Externalism is therefore an extremely important thesis. Not only does it, at least as Putnam believes, have serious consequences for what we can know empirically, but it appears to undermine the empiricist's conception of analytic truth.

Criticism of Semantic Externalism

Is Putnam's Semantic Externalism a tenable doctrine? Is it well supported by the considerations Putnam offered in its defense? This last question is obviously weaker than the first, for considerations other than the ones Putnam used may support it more strongly than his did.[45] But it is easier to answer this weaker question, and answering it may make it unnecessary to answer the stronger one.

The first thing to say about Putnam's defense of his semantic externalism is that he provides no clear account of the connection that he thinks is necessary for genuine reference. In one passage he appears to say that a genuinely referring predicate must be associated with "language entry rules" or "language exit rules":

---

[44] Putnam, p. 19.
[45] I discussed Tyler Burge's version of the doctrine in the penultimate section of chapter four.

> There are "language entry rules'' which take us from experiences of apples to such utterances as "I see an apple," and "language exit rules" which take us from decisions expressed in linguistic form ("I am going to buy some apples") to actions other than speaking. Lacking either language entry rules or language exist rules, there is no reason to regard the conversation of the machine…as more than syntactic play (p. 11).

But this claim is far too strong. Not every meaningful predicate is what a philosopher of science would call an observation term. We can surely talk about electrons, photons, and a host of other things without possessing language entry rules or language exit rules that feature the relevant predicates.

It is conceivable that in speaking of language entry and language exit rules Putnam meant to assert something far weaker—namely, that reference is possible only in a language containing basic predicates that are associated with such rules. This weaker thesis recalls the old doctrine of complex ideas, which I mentioned earlier in connection with Hume. According to this doctrine, some ideas are simple while others are complex. Complex ideas are built up from simple ones; we construct some of them and others arise from our interactions with complex objects. The ideas of a mermaid and a centaur are human creations; the idea of a dog or giraffe was no doubt originally generated in some human beings by the experience of perceiving such an animal. Other human beings got the idea from parents and friends who explained what these animals are like.

Do the words "mermaid" and "centaur" refer to anything? Given the sense in which Putnam uses the word "refer" in *Reason, Truth, and History,* we would have to say no. Putnam stipulates that he uses "refer" to stand for a relation that holds between a word, symbol, or idea and something that actually exists (p. 1, note). These words are perfectly understandable, however; we know what features a thing would have to possess to be a mermaid or centaur. Since predicates are general terms that purport to refer to many different things, the sort of reference they have is what empiricists used to call "multiple denotation."[46] As far as the word "mermaid" is concerned, this kind of reference is clarified by the formula:

M1 ∀x("mermaid" refers to x iff x is a mermaid).

A more revealing statement about the reference of "mermaid" is the following:

M2 ∀x("mermaid" refers to x iff x is like a woman from head to waist and a fish from waist to tail.)

M2 specifies a reference condition for "mermaid," and if a predicate of some language or conceptual system is associated with such a condition, we can say that it has a referential use even though it may lack an actual referent.

A weaker, more plausible claim that an empiricist might want to make about meaningful predicates is that they have a referential use only when they are associated with a reference condition that is specifiable by means of predicates that are themselves directly or indirectly attached to existing objects. The attachment to existing objects that these predicates have could be explained further by mentioning language entry rules, which a subject conforms to in making observations. The weaker claim I have been describing is not precise, but it is no more indefinite than Putnam's remarks about a predicate's causal connection to its referents. It does, however, raise a problem that is pertinent to the limits of acceptable reference.

---

[46] See Martin (1958), ch. 4.

The problem concerns the specificity of the relevant reference condition. If the condition is supposed to furnish necessary and sufficient conditions for a successful reference, it amounts to a definition of referential meaning in observational terms. This is evident from the fact that a language-entry transition conforming to a language entry rule is a propositional response to an experiential stimulus, an example of which would be thinking "That's red" when one has an appropriate red-sensing experience. But a definition of referential meaning in observational terms is tantamount to a positivist's conception of referential meaning. If the general condition that must be satisfied for acceptable reference involves a looser connection with language entry rules—one loose enough to permit reference to unobservable entities—it may then be perfectly acceptable, but it will not support Putnam's semantic externalism, for referring terms will not themselves have to be attached to anything that can prompt a language-entry response. Only a weak condition impresses me as realistic. We can meaningfully refer to leptons and quarks, which are in no way observable; but if we can do this, BIVs should be able to refer to brains, vats, and distant objects.

If we recall the basic structure of Putnam's argument for his semantic externalism, we can see that the case he made was exceedingly weak. He began by describing Turing's imitation game, which was intended to provide a test for answering the question, "Could a computing machine that successfully performed a certain imitative task be reasonably considered conscious?" He then asked if an analogous imitative task could show that a machine actually referred to something. He answered no, saying:

> What we have is a device for producing sentences in response to sentences. But none of these sentences is at all connected to the real world (p. 10).

He did not pose the general question "How must a language or discourse be connected to the world if words occurring in it are to refer to things in the world?' but he did say that unlike sentences that the machine might produce,

> Our talk of apples and fields is intimately connected with our *nonverbal* transactions with apples and fields. There are "language entry rules" which take us from experiences of apples to such utterances as "I see an apple" and "language exit rules" which take us from decisions expressed in linguistic form ("I am going to buy some apples") to actions other than speaking. Lacking either language entry rules or language exit rules, there is no reason to regard the conversation of the machine…as more than linguistic play (p. 11).

Putnam's last sentence here (on a plausible reading) is pretty clearly true, but it does not imply that every word that refers to something is associated with language entry or language exit rules. No doubt some rules of this kind are needed if the words of a language or discourse are actually *applied* to objects in the world, but Putnam does nothing to show that all referring words require such rules. Until he shows this, his case for semantic externalism is basically unsupported.

The idea that many referring terms are not associated with language entry rules is actually required for important claims Putnam makes about substances such as water. He makes these claims in slightly different ways in different essays. According to one statement, the referent of the word "water"[47] is identified by means of

[47] The word "water" is here functioning as a name, not a predicate; it names what I would call a natural substance rather than, as Putnam does, a natural "kind." As I see it, kinds are abstractions, not concrete realities.

paradigmatic samples whose chemical composition is ascertained by experts in chemical analysis. (He says the natures of other natural kinds are ascertained or identified by experts with other specialties: botanists are the experts to whom one appeals for information about the nature of plants, for instance.) Chemists tell us that water is H_2O; a substance on another planet that is superficially similar to our water but is not H_2O would not be water. According to another statement, paradigm samples of water are identifiable as such because of their appearance and the functional role of similarly appearing stuff in our world. Water is a transparent liquid that quenches thirst and makes plants grow; it falls from the sky as rain, fills lakes and ponds, and so on. Experts assure us that the substance in our world having these features and playing this role is H_2O. Since water *is* this substance—since it *is* H_2O—nothing could be water that is not H_2O. It is a necessary truth, one known empirically, that water is this chemical substance.[48]

If we ask what H_2O is, we will not want to be told "It is water." We will want information about the chemical formula. We will want to hear some story about hydrogen, oxygen, and the way these elements are related in H_2O molecules. But when experts start talking about hydrogen and oxygen atoms, they will be talking about entities that we cannot observe or interact with individually. There will be no language entry rules and language exit rules associated with the terms "hydrogen "atom and "oxygen atom". Putnam was fully aware of this fact. So he could not himself endorse the idea that general terms can refer to objects only if they come within the scope of some language entry or language exit rule.

To make a plausible case for the view that BIVs cannot refer to brains, vats, or any other object of what we normally consider the world, one will have to resort to considerations that Putnam did not identify. Is there anything about the assumed experience of such beings that would preclude their thinking about themselves, their tank, and the world outside their tank? If so, what is it?

I frankly cannot identify such a thing. Putnam is the only post-positivist philosopher I am aware of who has officially denied that beings as intelligent and as susceptible to empirical stimulation as we are cannot refer to what we can refer to, and the reasons for his denial are clearly unsuccessful. Lest the reader suppose that the patent difference between bodiless BIVs and our mobile selves must render their

---

[48] Although Putnam's semantic externalism is better suited to names than to predicates, his claims about water in this paragraph are obviously highly idealized and add little support to his externalist conclusions. No one supposes that a homogeneous substance actually fills all our lakes, ponds, and streams or that the liquids in those different geographical sites are chemically identical. Although we have very good reason to believe that any water we drink, swim in, or sail on consists largely of H_2O, our normal means of identifying a sample of water does not depend on this belief or on any other chemical lore. A chemist can tell us what proportion of a given liquid is H_2O or what other compounds it contains, but the decision to apply the label "water" to the liquid in the Cuyahoga river (which once caught fire), the Campus Pond at my university (which is often black and murky owing to the presence of thousands of migrating aquatic birds), the Dead Sea (which is heavily saline), or a highly diluted gallon of what was once Chardonnay wine, will not depend on such a person's decision. In fact, if our acid rain began to contain substantial amounts of the chemicals making up the XYZ liquid that fills the rivers and ponds of Putnam's Twin Earth without any significant effects on its ability to quench the thirst of animals or contribute to the growth of familiar plants, ordinary people would call it "water" without hesitation and continue to do so if, owing to some extraordinary natural change, it became pure XYZ. These and comparable other facts make it evident, I believe, that a meaningful reference to water does not depend, conceptually or semantically, on any set proportion of actual H_2O in the liquid a normal person is thinking of. A person with a smattering of chemistry might, of course, conceive of water as H_2O, but this conception would be anomalous in practice, for no water most persons have ever drunk is close to being pure H_2O. Good drinking water is heavily dependent on its mineral content. For a penetrating discussion of the relation between ordinary kinds such as water and molecular compounds such as H2O, one basically in agreement with what I say here, see LaPorte (2003), ch. 4.

references fundamentally different from ours, I should emphasize that the similarities between their kind and our kind are in some ways just as great as—and possibly even greater than—the differences. In fact, their intelligence and mental agility is supposed to be the same as ours, and their sensory input and conscious output—their sensory experiences and their awareness of what they are doing—are supposed to be "qualitatively" identical to ours. The differences between their experiences and thoughts and ours are limited to the way both are connected to external things and, consequently, to their supposed referential features. BIVs therefore have the ostensible experience of communicating with others and receiving responses from them; they have the experience of being members of a social community; and they ostensibly learn from others and provide instruction in return. But if we, from empirically identical experiential inputs and outputs, can develop a language that permits reference to trees and meadows, there is no apparent reason (other than the untenable ones Putnam provides) for thinking that the BIVs could not do exactly the same.[49]

Later in *Language, Truth, and History,* Putnam criticizes the very notion of objective reference, arguing that it belongs to the perspective of "metaphysical realism," which he rejects.[50] He calls his own perspective "internalism" and says that for those accepting this perspective reference makes sense only "within a conceptual scheme." We "cut up the world into objects when we introduce one or another scheme of description," he says, and because "the objects signs are alike *internal* to the scheme of description," it is "trivial to say what any word refers to *within* the language the word refers to":

> What does "rabbit" refer to? Why, to rabbits, of course! What does "extraterrestrial" refer to? To extraterrestrials (if there are any)…. For me [Putnam says] there is little to say about what reference is within a conceptual system other than these tautologies. The idea that a causal connection is necessary is refuted by the fact that "extraterrestrial" certainly refers to extraterrestrials whether we have ever causally interacted with any terrestrials or not![51]

In saying this, Putnam dismisses, virtually without argument, the skeptical problem that he attempted to dispel by a serious argument in the first part of his book.

The minimal argument Putnam gives for this later position concerns the perspective of the person framing the BIV hypothesis. Certainly no BIV would advance this hypothesis, he says; and if a non-BIV were to advance it, "the world would not be one in which *all* sentient beings were Brains in a Vat." So, he concludes, the BIV hypothesis "presupposes from the outset a God's Eye view of truth, or, more accurately, a No Eye view of truth—truth as independent of observers altogether."[52] And this is incompatible with his internalist perspective. A metaphysical realist might attempt to pose the problem, but his or her assumptions about reference and truth would render the attempt futile, since a BIV could not, on those assumptions, entertain the hypothesis at all.

This way of disposing of the BIV hypothesis is far too simple. The question of how we can know that certain sorts of unobservables exist is theoretically significant, and the perspective of the being who advances the BIV hypothesis is not sufficient to

---

[49] A detailed explanation, based on neural inputs and brain physiology rather than external objects, of how human beings can form a "mental" representation of themselves and their environment can be found in Trehub (1991).

[50] Putnam's metaphysical realism is a peculiar doctrine that no actual philosopher, to my knowledge, ever espoused. I criticize it in Aune (1985), pp. 126-28. William Lycan caricatures it in a wonderfully amusing way in Lycan (1988), p. 190.

[51] Putnam (1981), p. 52.

[52] *Ibid*, p. 50.

refute it. Any sane person believes that he or she inhabits an objective world of animals and things, but Putnam's story of brains in a vat is coherent and describes a conceptual possibility, which philosophers normally find interesting to think about. If, as we believe, we can think about objects we cannot actually observe—if we can think about electrons and photons no less than prime numbers and algebraic functions—the same should be true of BIVs: they should also be able to think about things that are not, for them, observable. Contemplating such a possibility does not require some philosophically objectionable "perspective."

A Skeptical Problem Restated

When we contemplate the possibility of BIVs thinking about objects they cannot actually observe, an old epistemic problem arises again. It arises from the similarities I emphasized between our thoughts and experiences and those of the BIVs in Putnam's story. Although we certainly believe that we experience shoes and ships external to us in space, we know that there must be a flow of information from those objects to us, and the last part of this flow is qualitatively the same as what a deluded BIV is supposed to experience when it thinks it is perceiving a shoe or a ship. Another similarity is present in what Putnam calls language-exit transitions. When we implement an intention to reach for an apple, we have the experience of reaching for an apple—and a qualitatively identical experience would occur in a BIV according to Putnam's story. The sequence of events beginning with real external objects and our perception of them as well as the sequence of events beginning with our motivating intentions and ending with our overt actions contain segments consisting of conscious experiences; and these intermediate segments may be presumed to be empirically the same in us and in the BIVs. How, then, can we *know* that we are actually physically different from the BIVs?

As I observed in chapter one, we normally brush aside the possibility that we might be anything like BIVs. The idea is too far-fetched to be taken seriously in everyday life. But philosophical reflection, at least epistemological reflection, is not a staple of everyday life. In everyday life we say we know all sorts of things whose truth we ascertain or surmise only by means of presumptions that sometimes fail. One such presumption, the defeasible presumption I mentioned in chapter one when I discussed the example of the phony barn perceived through the window of a train, is that we actually see what we seem to be seeing in the light of day. Lewis mentioned other such presumptions in giving his Rule of Reliability and his Permissive Rules of Method.[53] Everyday ascriptions of knowledge are based on these presumptions, but they are defeasible and always questioned by philosophers in search of certainty. They are interested in proof, and presented with Putnam's story, they will want to know if there is anyway of proving that we are not BIVs. Putnam purported to provide such a proof in advancing his semantic externalism argument. His argument failed. Is an alternative available?

For an empiricist, a proof is out of the question here. It is even out of the question for a philosopher like Hume, who, as I mentioned in chapter one, spoke of a kind of proof in principle applicable to empirical propositions: it is provided by "such arguments from experience as leave no room for doubt or opposition." A typical empiricist—and this includes Hume—would contend that even in this weak sense a proof is not possible for the proposition that what I called mediating experiences are in fact connected to external objects. Why not? Because the connection is causal and purely contingent. Such connections cannot be ascertained a priori. They can

[53] See chapter one, p. 12.

be "known" only by a posteriori inference, and this kind of inference yields probabilities rather than certainties.

As it happens, there are unresolved problems about the logical structure and rational acceptability of the forms of inference by which such connections can presumably be ascertained. These forms of inference are, in fact, needed to provide rational support for a significant variety of familiar beliefs—for instance, those about the remote past, the experiences of other people, and unobservable objects such as electrons. Beliefs about these things have always been problematic for empiricists. I shall discuss these problems, together with the subject of memory, one of the empiricists' three sources of empirical knowledge, in the chapter to follow.

Chapter 6

MEMORY AND A POSTERIORI INFERENCE

Empiricists have always regarded observation and memory as the fundamental input for empirical knowledge; they say that further knowledge is obtained mainly by a posteriori inference. In the last chapter I discussed the sort of knowledge that can be obtained by observation. After just a few pages it became obvious that observational knowledge is very closely related to inferential knowledge; in fact, much of what we think we know by observation actually requires some kind of a posteriori inference. It will turn out that memory knowledge is basically the same as observational knowledge in this respect. A posteriori inference is therefore fundamental to most empirical knowledge. I shall be principally concerned with salient varieties of a posteriori inference in this chapter, but I begin with a brief discussion of memory and eventually conclude with an empiricist defense of the presumptions on which our everyday claims about the external world are ultimately based.

Memory as A Source of Knowledge

Although empiricists have always regarded memory as an indispensable source of our empirical knowledge, they have generally praised memory with a slightly bad conscience. The problem with memory is that it purports to tell us something about occurrences that do not exist when the memory occurs. A recollection occurs in the present; the occurrence it refers to took place in the past. But the past is over and done with, and the same is true of the occurrences we ostensibly remember. Any correspondence between a current memory and a past occurrence cannot therefore be ascertained by direct inspection. Since we cannot infer such a correspondence by a priori reasoning, there is no inconsistency in the supposition that the world came into existence just a moment ago, complete with the recollections we now have. This possibility cannot apparently be ruled out by a posteriori reasoning, at least if that reasoning is the sort of thing empiricists have taken it to be: a matter of generalizing from experience. If we never directly experience a connection between present impression and past occurrence, we have absolutely no basis for any such generalization. By what right can we trust memory if we have no way of proving that it is ever correct?

I have just rehearsed the classic case for a skeptical view of memory. At first sight, it is a very impressive case, but reflection shows that it is highly exaggerated. Do we not sometimes have direct perceptual access to part of the past? If we can perceive things in motion, we must have this access. When someone smiles, waves a hand, or throws a ball, the action takes time, and the first part of the action is over (a past occurrence) when the last part occurs. The perception of any movement is thus attached to a temporal interval that includes past, present, and future. The past and future occurrences are past and future in a relative sense: in an occurrence with three distinguishable segments A, B, and C, the segment B is a past occurrence in relation to C but a future occurrence in relation to A. This relative futurity is a genuine kind of futurity because when A begins to occur, B has not yet taken place. When B does take place, A has taken place.

Saint Augustine, who was seriously perplexed about time, had a very different view of the present. As he explained it,

> If an instant of time be conceived which cannot be divided into the smallest particles of moments, this only is it which may be called present…. For if it be, it is divided into past and future. The present has no space.[1]

If Augustine was right here, the present is a timeless moment. But we are not conscious of a timeless moment when we make observations. The world we perceive or otherwise experience is always in motion: it is consciously changing, and we experience it as changing. The idea of a "spaceless" present is created by a process of abstraction, by thinking away the terminal elements of an experienced moment. The idea thus created is comparable to the idea of an imperceptible triangle. Our knowledge of it results from inference, not observation. The moments we observe contain relative pasts and futures as well as extended nows.

As strange and possibly puzzling as the idea of an extended present may be, we have as much reason to believe we perceive extended moments containing changing things as we have to believe that we perceive anything at all. Of course, the amount of the past included in what we can perceive is relatively little: I can perceive a smile, but I probably do not perceive the whole of a forward pass in football. I watch a quarterback move his arm forward, I see the release, and I watch the ball move down the field. I suppose it is more accurate to say that I observe a series of movements. When I observe the later movements, I am recalling the earlier ones. I know what these are like because I remember observing them.

Philosophers of an earlier time devoted a lot of attention to the phenomenology—the qualitative aspects—of memory experience.[2] Thinking about this aspect of memory is no longer fashionable in philosophy; perhaps it is now supposed that the experience of remembering may be different in different people. But the qualitative character of remembering does not seem important for epistemology anyway. What is important is the truth or probable truth of the claims people make about what they remember. People who say they vividly remember certain things are generally very confident in the truth of what they say; those who declare that their memories are dim are usually less confident; and those with memories of intermediate vividness have intermediate confidence in what they say they recall. Of course, some people are naturally more cautious or more conscientious than others; some are even more interested in calling attention to themselves or in telling a good story than in being right. The timid claims of some are therefore sometimes more trustworthy than confident claims of others.

When people claim to recall things they once experienced, their recollections can often be supported or criticized by other records of the past—by diaries, letters, photographs, films, and the like. Because of such things, we do not have to rely entirely on a person's words for our picture of what actually happened. Yet words are centrally important for many past occurrences. What Tom promised Ted or what Sally told her students on Friday could never be known in any other way. To decide whether this or that person's recollection is correct, we must in fact consider the variety of factors that are pertinent to the assessment of an observation report. Since people ostensibly remember what they saw, heard, tasted, or learned in some way, the truth of what they remember depends crucially on the truth of what they think they perceived or otherwise learned. In assessing the probable truth of memory impressions or reports, we must therefore take into account the considerations pertinent to evaluating an observation report in addition to those specifically applicable to the reliability of a subject's memory and the motives he or she may have for embel-

---

[1] Augustine (1952), p. 226.
[2] See Russell (1948), pp. 226-232.

lishing or even falsifying a true recollection. If the truth of a certain memory claim or the occurrence of an event a person was in a position to recall is very important, as it commonly is in a legal proceeding, we might insist on having the subject cross-examined by a competent lawyer. Not only will our existing evidence be tested by the cross-examination, but further evidence will also be obtained. As every reader of mysteries knows, the process of discovering what actually happened on this or that occasion can be extremely complicated.

In spite of the complications that I have just emphasized, most memory claims might be described as past-tense observation claims: they may differ from a typical observation claim by no more than "I see Spot run" differs from "I saw Spot run." I emphasized the variety of considerations pertinent to evaluating an observation claim in chapter five; the complications pertinent to assessing a typical memory claim may be no greater than those pertinent to its present-tensed cousin. In fact, if a person is seriously questioned about what he or she now observes, the duration of the questioning may easily convert the target of the investigation into a memory claim. What is logically special about a memory claim—what makes it deserving of separate treatment—is that the inferences properly supporting its truth or probability are essentially backward looking. A fact about the past is inferred from facts about the present.

What sort of inference is capable of providing this kind of support? Hume considered it experimental. As I explained in chapter five, the reasoning Hume called "experimental" is causal inference; it consists in inferring one fact from another by means of a causal principle obtained from experience. If we represent a certain causal principle by "As cause Bs," we can identify two associated forms of experimental inference. One infers Bs from As, or effects from causes; another infers As from Bs, or causes from effects. When Hume treats particular causal inferences in detail, as he does in his Dialogues Concerning Natural Religion, he emphasizes that the causes and effects appropriate to a given causal principle have to be identified very carefully; but his basic idea is that "all experimental reasonings are founded on the supposition that similar causes prove similar effects and that similar effects prove similar causes."[3] If Hume is right, the inferences supporting the truth of memory claims are experimental inferences of his second kind: a past cause is inferred for present effects.

The cause that is inferred by such an inference is the occurrence that is ostensibly remembered. What effects provide a proper basis for such an inference? If the emphasis is squarely on memory and not other effects of past occurrences, the relevant effect is probably the subject's memory experience or memory belief. Suppose I have the experience of ostensibly remembering (or seeming to remember) parking my car a half hour ago in section C4 of the parking lot outside the restaurant where I am now having lunch. If I can assume that this kind of experience is probably caused by the sort of actual occurrence it seems to represent, I can conclude that I (probably) did park my car a half hour ago in section C4 of that parking lot. In Hume's view, all the inferences that provide empirical support for the truth or probable truth of memory claims are of this causal kind. If a given claim receives additional support from a note written in a diary, the note must be viewed, if Hume is right, as an indirect effect of the occurrence it describes.

I noted in chapter five that Hume thought the causal principles used in experimental inferences resulted from experience but were not themselves inferred from any premises at all. Later empiricists generally disagreed with Hume on this last point; the consensus was, and possibly still is, that causal principles are obtained by induction, an inferential process also known as inductive generalization or enu-

[3] Quoted in Kemp Smith (1948), p. 147.

merative induction. The skeptical view of memory that I described early in this chapter was based on the idea that inductive inferences of this kind are rationally unjustifiable. If inductive inferences can be justified, memory claims can no doubt often be justified in the way Hume thought—by causal reasoning. But there are many problems with induction. Now is a good time to see what they are.

What is Induction?

One of the problems about induction is how the rule should be formulated. One way of presenting this problem is to show the defects of a commonly offered formulation. The one I shall begin with is given by William Lycan (1988), and it is similar to a formulation used by Laurence BonJour. Both writers accompany their formulations with qualifying remarks, Lycan's identifying fallacious applications of the rule. His rule is this:

> n% of all the observed Xs have been F.
> Therefore [probably] roughly n% of all Xs are F.[4]

The fallacious applications he has in mind, which his formulation does not itself rule out, occur when the number of observed Xs is too low to represent a "sufficient" sample or when the observed Xs constitute a biased one. If your evidence class—that is, the class of observed Xs—were very small, you would normally have a very poor basis for claiming that approximately n% of all Xs have the property F.[5] Similarly, if the Xs you observe were not selected in some impartial or fair way, you would normally have a very poor basis for making a comparable claim.

In view of these fallacies, it is important to look for a rule that disallows them. Consider the following:

> n% of all the observed Xs have been F.
> A representative example of Xs have been observed.
> Therefore, [probably] n% of all Xs are F.

This formulation would no doubt disallow generalizations from insufficient and biased samples, but to apply it, we would have to know how we are to identify a representative sample. Suppose we are told that a sample of Xs is representative of a larger reference class with respect to the frequency of having F just when the percentage of Xs having F in the evidence class is approximately the same as the percentage of Xs having F in the reference class. If this is what we are to understand by a representative sample of Xs having F, the revised inference schema would be deductively valid: the corresponding conditional statement would be analytically true. This would give us an unquestionably valid form of "inductive" inference, but we would have no way of knowing when a particular evidence class is representative in the specified sense.

In Human Knowledge: Its Scope and Limits, the only one of his many books in which he seriously discussed inductive inference, Bertrand Russell offered the following as an inductive rule:

> Given a number n of α's which have been found to be βs, and no α which has been found to be not a β, then the two statements: (a) "the next α will be β",

[4] Lycan (1988), p.179.
[5] The word "normally" appears in this and the following sentence for a reason that will become evident as the discussion proceeds.

> (b) "all α's are β's", both have a probability which increases as n increases, and approaches certainty as a limit as n approaches infinity.[6]

To apply this rule we do not have to know whether we have a representative sample of αs and βs, but the rule will give us little help if we are interested in drawing a conclusion about the percentage of native-born Norwegians having blond hair. To draw a conclusion about a reference class that, although finite, is too large to examine as a whole, we shall need some way of estimating the size of an acceptable evidence class and of identifying an impartial way of selecting its members. Unfortunately, no general description of how these tasks may be accomplished appears to be available.

A possible reason for the dearth of general descriptions is that dramatically different sample sizes and methods of selection appear to be acceptable in different cases. Consider the way new models of automobiles are evaluated each year by Consumer Reports. Normally, just one example of a given model is examined, and the example is obtained merely by buying it from some randomly chosen dealer without disclosing the actual identity of the buyer. Although one might initially suppose that a single example is far too small to be an acceptable evidence class, reflection shows that a single example is almost certain to be representative of the model to be sold with respect to traits deemed important for the entire class. The reason for this is that automobiles are mass-produced objects subject to standard quality controls. Some manufacturers produce more reliable products than others do, but a given manufacturer is apt to produce instances of a particular model in the same way using basically the same materials. Anomalies occur, of course, but one instance can be expected to be substantially similar to any other instance of the same model, particularly if the dealer selling it has no reason to suppose that the buyer will use a particular instance in a way that will compromise future sales.

The acceptability of the sample size and the method of selecting instances in this last case obviously depend on background information about automobiles and the way they are produced. Here the acceptability of one inductive inference appears to depend on the acceptability of others. The question therefore arises, "Is there is a basic rule for rationally compelling enumerative induction whose application does not require background knowledge of this kind?" As far as I know, the answer is no. Hume, assuming no relevant prior knowledge of the objects of an empirical generalization, argued that an evidence class, no matter how selected and how extensive it may be, provides no rational basis for the conclusion that the objects of the reference class, many of whose members may exist in the distant past or the remote future, are at all similar (in the relevant ways) to those already examined. We naturally expect them to be similar—"we expect the future to be [relevantly] like the past", he said—but this expectation has no basis in reason or any operation of the understanding. It is purely instinctive. There is no inconsistency supposing that the future will be unlike the past in relevant respects; and any a posteriori reason that could be offered to dispute this would be based on the same supposition and thus beg the question at issue.[7]

Induction: Arguments Pro and Con

Laurence BonJour recently countered Hume's criticism with an a priori argument, one featuring a form of inference that Hume did not consider. BonJour's a priori argument, which is of course a defense of enumerative induction, is noteworthy for two basic reasons. It includes novel qualifications to the inductive rule designed to avoid

---

[6] Russell (1948), p. 419.

[7] See Hume, *Enquiry,* Sect. IV.

objections raised only in recent times, and it relies on an additional form of inference that is now fashionable with philosophers and deserving of critical attention.

BonJour's argument applies to what can only be called a very incomplete formulation of an inductive rule. Initially, he identifies the sort of situation in which an inductive inference could (as he sees it) be reasonably made. The situation would involve "a large number of observed instances" of something A, a fraction m/n of which have "some logically independent observable property" B. The locations and times of observation, the identity of the observers, the conditions of observation, and any further pertinent background circumstances must be varied "to a substantial degree" and there must be no relevant background information available concerning either the incidence of Bs in the class of As or the connection, if any, between being A and being B."[8] If these conditions are met and the observed proportion of As that are Bs "converges over time to the fraction m/n and thereafter remains at least approximately constant as significant numbers of new observations come in," then the conclusion of the argument is likely to be true (p. 207).

As I noted, BonJour attaches some important qualifications to the inference he describes here, but his a priori argument for its acceptability—for the fact that its conclusion is probably true when its premises are true—is that the truth of the conclusion provides the best explanation for the data that the premises describe. The crucial data here pertain to the "convergence and constancy of the observed proportion," and although it is possible, he says, that this proportion is a matter of chance, it is highly likely that the observed proportion is an accurate reflection of an objective regularity (p. 209). Such a regularity cannot be a mere constant conjunction, as Hume thought; to provide the required explanation it must be a "metaphysically robust" regularity, involving a necessary connection (p. 215) or a "substantial propensity to persist into the future" (p. 214).

The qualifications BonJour adds to his account apply to two kinds of counterexamples independently discovered by Bertrand Russell and Nelson Goodman in the late 1940s. Both philosophers observed that any objects chosen as the basis for an inductive generalization possess some features that support objectionable generalizations—generalizations that are either patently false or incompatible with other generalizations that are equally well supported by the available evidence.

Russell's examples showed his usual wit. One was based on the well-known belief that Immanuel Kant had never been more than ten miles from his hometown of Königsberg.[9] If Kant had been interested in drawing inductive conclusions about sheep, one property that he might have observed in every sheep he examined is that of being within ten miles of Königsberg. To get as large a sample as possible, he could have devoted years to the task of observing sheep and, to make his selection as unbiased as possible, he might have observed them in fields, in barns, on houses, and possibly even in ponds. A generalization supported by his observations would have been the patently false "All sheep are within ten miles of Königsberg." Other properties possessed by every sheep he might have observed are being observed by Immanuel Kant, being observed by someone, living in Germany, living in Europe, and being outside of Italy. Obviously, this list could be extended indefinitely.

Goodman's examples featured contrived predicates such as "grue," the latter applying to an object, Goodman stipulated, just when it is either green and examined before a distantly future time t or blue and not so examined. Goodman argued that if we are examining emeralds for color and find that they are invariably green, we can use the inductive principle to draw two incompatible conclusions, neither of

[8] BonJour, pp. 188f.

[9] Russell used this last property to illustrate the "shaky" character of induction by simple enumeration in Russell (1951), p. 126.

which is better supported than the other. They are "All emeralds are green" and "All emeralds are grue." These conclusions are incompatible because they disagree about the color of emeralds not examined before t. According to one, they are green; according to the other, they are blue. These conclusions are equally well supported because every emerald we examine will inevitably be examined before t and thus, if it is green, count as grue. Since the time t may be placed arbitrarily far in the future, we cannot avoid the difficulty by waiting to see how emeralds look when t arrives. We are, in fact, faced with a general problem arising from the inductive principle itself. It permits us to draw incompatible conclusions from the same body of data. This defect is illustrated by the hypothesis featuring the word "grue," but it is common to countless other hypotheses. To avoid it, the inductive principle needs significant qualification.

As you might expect, one of BonJour's qualifications requires that the predicates used in an acceptable inductive inference do not include the contrived sort exemplified by "grue" (p. 189, note 2). Unfortunately, this qualification is not actually effective in avoiding the problem Goodman raised. The choice of predicates is in fact irrelevant to the issue. Instead of using the predicate "grue," Goodman could raise his problem simply by speaking of things that are either green before the time t or blue afterwards. If every emerald we examine is green, every emerald we examine is either green before t or blue after t. Because of this, the emeralds support the hypothesis H2, that all emeralds are either green before t or blue afterwards, just as strongly as they support the hypothesis H1, that all emeralds are green. But H2 conflicts with H1 in regard to the color emeralds have after t. This problem is evidently not avoided by the inductive rule that BonJour attempts to justify. The problem does not arise, incidentally, from the disjunctive character of one of the hypotheses. It can arise just as easily from curve fitting problems.[10]

When Goodman introduced predicates such as "grue," his aim was to call attention to what he called a new problem of induction. This new problem was essentially the same as the one Russell raised: the familiar inductive rule needs serious qualification if it is to be acceptable. Russell did not suggest a qualification for the standard rule; Goodman did. Oddly enough, the qualification Goodman offered bears an interesting similarity to the principal qualification BonJour provided. Somewhat like BonJour, Goodman ruled out troublesome hypotheses—the one's containing words with the same extension as predicates such as "grue" [11]—on the ground that they are not "lawlike."[12] He distinguished lawlike from non-lawlike hypothesis by reference to a property that he called "entrenchment," which BonJour did not come close to mentioning, but he did imply that the hypotheses strongly confirmed by their instances could be regarded as laws, or statements of laws. This recalls BonJour's claim that regularities inferable by enumerative induction must be "metaphysically robust" regularities, involving necessary connections or "substantial propensities to persist into the future."

Induction and Laws

The idea that the generalizations reasonably inferred from data samples are or must be laws is so implausible that it is hard to take seriously. People who conduct public opinion polls draw general conclusions from their samples, but they rarely if ever suppose that their conclusions hold true eternally or even far into the future. The same is true of conclusions about the effects of advertisements, the fear of epi-

[10] See Grunstra (1969), pp. 102-106.
[11] Predicates P1 and P2 have the same extension just when they apply to the same objects.
[12] Goodman (1965), p. 73.

demics, or any of the thousands of topics that are investigated by statistical methods every year. There is nothing "lawlike" about the conclusion that the U.S. President's approval rating among voters is ten per cent less today than it was two months ago, even though this conclusion was inferred from samples taken all over the country.

Some of the generalizations inferred from experimental data might, I suppose, be considered "laws," although the very idea of a scientific law is less widely accepted these days than it used to be.[13] But there is no agreement among statisticians that inferred laws are generally more secure than short-term generalizations about public opinion.[14] If this is right, then if enumerative induction deserves to be regarded as an acceptable form of a posteriori inference, BonJour's a priori justification at best applies only to a limited class of these inferences—and not to a favored class whose members are used with greater confidence than the others. The qualifications he places on the kinds he defends do rule out some of the counterexamples Russell constructed, but they do not succeed against Goodman's counterexamples, which do not really depend on special predicates, nor do they succeed against the full range of counterexamples Russell had in mind, which are essentially the same as Goodman's.[15]

Like BonJour, Goodman wanted to disallow the "bent" hypotheses he discussed as well-confirmed examples of inductive conclusions, and he did so, I said, by claiming that they are not lawlike and so not confirmable by their instances. But Goodman's solution to his new riddle of induction is arguably too restrictive even for the case of scientific "laws". Citing specific scientific theories, Rosenkrantz (1981) persuasively argued that scientific advances often result in hypotheses that are more "bent" than the ones they supersede. The grue hypothesis, he said, in fact "belongs to a class of hypotheses that are not only scientifically quite respectable but are the very ones whose introduction so often marks the breakthroughs we are wont to label 'scientific revolutions'" (p. 7.1, 4). The price of adopting Goodman's "entrenchment" solution to the new riddle, Rosenkrantz contends, is much too great to tolerate.

An enormous literature has grown up around Goodman's new riddle and his proposed solution to it, and Rosenkrantz's criticism, as impressive as I find it, is no doubt not the last word on the matter. Specialists in the history and methodology of science can speak to it far more effectively than I can. But BonJour's a priori defense of induction has another feature that raises important issues of a different kind. It is based on what is now known as an inference to the best explanation, a form of inference that is widely regarded as a posteriori rather than a priori. I will discuss the logical structure of this kind of inference a little later; right now I want to say something about BonJour's belief that the inference he employs in defending a schematic example of enumerative induction is a priori.

Judging by the steps he takes in constructing his argument, I think it is fair to say that BonJour's belief in this matter rests on two assumptions, which he thinks he knows to be true a priori:[16]

1. The best explanation (meaning "best explanatory account") that can be given for a body of data is most likely to be true.

2. The best explanation that can be given for the truth of a standard inductive premise is the straight inductive explanation, namely that the ob-

[13] See van Fraassen (1989), Part 1, pp. 15-128.
[14] See Phillips (1974), chapter 6, for an elementary discussion of probability densities.
[15] See Russell (1948), p. 422.
[16] See BonJour's principle (I-2) in BonJour (1998), p. 212.

> served proportion m/n reflects (within a reasonable degree of approximation) a corresponding objective regularity in the world.

In formulating these two assumptions I am ignoring some claims that BonJour makes in the course of his argument but that are actually not needed for it.[17] The "best explanation" that he identifies in (2) is clearly the best that can be given in his opinion; and the truth of (1) is something he thinks he can simply see to be true. The assumption I now want to say something about is (2).

BonJour thinks what he calls "the straight inductive explanation" (or SIE) is the best one for two reasons. First, he thinks the connection between A and B observed in the evidence class must be explainable by some law. And second, he thinks that any genuine law consistent with the evidence but requiring a divergence from the observed ratio m/n in a way that would falsify SIE would not really be possible. A genuine law requiring a divergence from m/n could be owing only to a further characteristic C, he says, one that affects the facts of observation itself, and this runs afoul of one of the qualifications he mentioned in describing acceptable induction in the first place.

These two reasons are idiosyncratic and certainly not convincing. As for the first, I can think of no tenable basis for supposing that a regularity observed to hold during some finite interval, however long, can be explainable only by a law. A more extensive regularity, one without temporal limits, will certainly do as well. An actual law is not needed, I should say, because explanation, pragmatic considerations aside, is prediction after the fact, and anything predictable by a law is equally predicable by a temporally unrestricted regularity: the modal character of a law, its supposed necessity, has no distinctive observable consequences. As for the second reason, an appropriate regularity between A and B can certainly be such that m/n of As are B in one spatial or temporal region but j/n of As are B elsewhere (j being significantly larger or smaller than m). The variation can simply be a matter of the way A is related to B; another characteristic is not needed to account for the divergence from m/n. The sort of "bent" hypotheses (or supposed laws) that Rosenkrantz cites in criticizing Goodman provide actual examples of such lawfully predicable divergences.[18]

If I am right about these last points, BonJour's attempted priori justification of enumerative induction does not succeed; it does not even overcome the arguments casting doubt on the idea that enumerative induction does not deserve to be considered an acceptable form of inference. But there is a further matter to be discussed, the acceptability of the form of inference that BonJour relied on in his attempted justification--namely, Inference to the Best Explanation or IBE. This form of inference is now widely accepted; in fact, some well-known writers—for example, William F. Lycan—regard IBE as the basic form of a posteriori inference. According to Lycan, enumerative induction can be reconstructed as a special case of IBE.[19]

---

[17] One assertion that does no work in his argument is that "it is highly likely that there is some explanation (other than mere coincidence or chance) for the convergence and constancy of the observed proportion…." (p. 208). This is obviously not needed if he can simply identify the best explanation.

[18] Rosenkrantz (1981) describes these hypotheses as positing "theoretically well-founded deviations from an overriding 'straight' hypothesis at extreme ranges of the relevant variables." His examples are taken from both special and general relativity. See Rosenkrantz, ch. 7, sect. 1. See also Rosenkrantz's lucid paper (1982), which deserves to be considered a classic on the philosophy of induction.

[19] Lycan (1988), 178-188. A more extensive discussion is given in Lipton (2004), but Lycan's description is adequate for the task at hand. Bas C. van Fraassen and others discuss critically the second edition of Lipton's book (2004) in van Fraassen (2006). See below, footnote 21.

Inference to the Best Explanation

To evaluate this form of inference, one must understand its logical structure. Lycan describes this structure as follows: [20]

1. F1,…,Fn are facts.
2. Hypothesis H, if true, would explain F1,…,Fn.
3. No available competing hypothesis would, if true, explain the Fi as well as H does.
4. Therefore, [probably] H is true.

There is no doubt that we often reason according to this pattern in everyday life, but it is actually very doubtful whether the practice is as commendable as Lycan supposes.21 The third premise should raise immediate doubts. What count as competing hypotheses in a particular case? If we are to apply the method, we must be able to survey all the available "competing" hypotheses. If a pertinent hypothesis is overlooked, we cannot be sure we have found the best explanatory account. The class of competing hypotheses must therefore be limited to those we can think of; they cannot comprise a class of ideal alternatives. But why should a reasonable philosopher suppose that the right explanation for the Fi is generally supplied by one of the hypotheses some actual person can think of? It is obvious that most of the facts we can explain today by the quantum theory or the theory of relativity could not be rightly explained by any hypothesis that Plato or Aristotle could even conceive of. Are we to suppose that we are bound to be in a better position to explain an arbitrary occurrence than they were? If a phenomenon is similar to others that we have successfully explained by accepted principles, we can approach it with a fund of knowledge that may assist us in identifying the likely explanatory factors. But if we lack this knowledge, the account that seems best to us might be wide of the mark and certainly not "probably true."

As it happens, real-life inferences to the best explanation are commonly fanciful and irresponsible. People not trained to weigh evidence confidently offer explanations in cases where they lack the information to provide any reasonable explanation at all.[22] A skeptical attitude does not seem to be natural to ordinary human beings. When a fanciful explanation is offered for some fact, judicious observers are apt to reject it immediately even if no alternative explanation is available; they do so because they doubt the proffered account is actually true. If they do accept an explanatory account, it is only because it is already significantly credible; in choosing it, they are generally convinced that it rather than some other acceptable principle is applicable to the facts in question. If no generally accepted principle seems to apply, they may speculate about a possible explanation, but they never, if they are judicious, actually accept an explanatory hypothesis as "probably true" (at least they ought not to do so) if they merely regard it as preferable to the other explanatory accounts that they can think of. They might regard it as providing a possible explanation that deserves to be kept in mind and tested further, but they would not accept it as "probably true" if it had nothing else in its favor.

---

[20] *Ibid*. p. 129. I alter Lycan's description in trivial ways compatible with his intent. He uses "explain" where I use "would explain if true," but he says he is using "explain" in the "nonsuccess" sense equivalent to "would explain if true."

[21] van Fraassen (1989) criticizes this form of inference in some detail. His objections are in general agreement with the objections I develop here. See also his contribution to the review symposium on the second edition of Lipton's *Inference to the Best Explanation* in van Fraassen (2006), pp. 344-352.

[22] Evidence for this is given in Tversky and Kahneman (1983).

About the only time a hypothesis is regarded as strongly supported by its evident success in explaining certain observed facts is when that hypothesis is antecedently probable, or already acceptable in a significant degree, and the facts are antecedently improbable, not predictable by other accepted principles. An example of a hypothesis supported this way was Einstein's general theory of relativity; photographs of fixed stars taken during a solar eclipse supplied the supporting facts.[23] Before the photos were taken, the theory was not regarded as sufficiently probable to be accepted, but the disjunction "Either Einstein's theory is true or Newton's theory of light and gravitation is true" was regarded as highly probable, and not entirely owing to the probability of the Newtonian disjunct. The facts were antecedently improbable, because the fixed stars had never been observed in the precise arrangement predictable by Einstein's theory and verifiable by the photographs. Unlike the hypothesis supported in this example, one with a low antecedent probability generally receives only weak support from the data predictable by means of it. To obtain nontrivial support for such a hypothesis, persistent testing will normally be needed. Those who favor the hypothetico-deductive method emphasize the importance of such testing, but it is not even suggested by Lycan's description of Inference to the Best Explanation.

The HD or hypothetico-deductive method is a kind of precursor to IBE.[24] According to it, a hypothesis is tested by deducing consequences from it in conjunction with auxiliary assumptions that are considered true or approximately so. If the consequences are verified, the hypothesis is confirmed in a positive degree; if they are refuted, the hypothesis is amended and tested again, or simply rejected. This kind of testing is supposed to be persistent, for a hypothesis is considered acceptable only if it stands up to a considerable amount of testing. In this respect, its confirmation is similar to what a hypothesis is thought to require from enumerative induction. Understood as I have described it, the HD method seems a bit simple-minded, because the testable consequences of one hypothesis can always be inferred from another one. (This is trivially true, because a new hypothesis can be made from an old one by adding on some qualification, humdrum or exotic.)[25] For this reason, the HD method is naturally modified along the lines of Inference to the Best Explanation. The aim now is to choose the best hypothesis from a family of alternatives. Testing is by prediction, as before, but now alternatives covering the same data must compete. The best hypothesis should ideally be the simplest, the most testable, and one that fits in best with background knowledge.

As I see it, the modified HD method, like Lycan's description of IBE, does not accord with reasonable scientific methodology. For one thing, it possesses one of the basic defects of IBE: the family of alternatives it features are just the alternatives someone can think of, and there is no a priori basis for thinking that one of these alternatives is apt to be true. For this, additional knowledge is needed. For another thing, the probability of the conclusion we can expect to infer from observed data is bound to admit of degrees: some conclusions are weakly supported at best; others are supported more strongly. The antecedent probability (or plausibility) of the other hypotheses should also be taken into account, and the same is true of the supposed facts that these hypotheses are supposed to predict or explain. Clearly, some hypotheses are more far-fetched than others, and the existence of antecedently unlikely facts predicted by a hypothesis will support it far more strongly than will those

---

[23] See Carnap (1956), pp. 158ff.

[24] See Aune (1970), pp. 167-182, and Earman (1983).

[25] The HD method does not require that each alternative hypothesis must be antecedently probable in a significant degree. A requirement of this kind is appropriate to the view of confirmation I discuss in the next section.

that are likely to occur anyway. A conception of experimental inference that does not accommodate these probabilistic considerations cannot be deemed satisfactory. Inferences conforming to the HD method or Lycan's pattern are generally dubious, I should say, because they ignore too much that is pertinent to the support of an acceptable hypothesis.[26]

Inferences Based on Bayes' Theorem

There is an alternative form of inference that does not possess the limitations I have just mentioned. It is based on the use of a simple theorem of probability theory, one known as Bayes' theorem. Ordinary people and even most scientists rarely employ this theorem in routine inferences, but then they rarely employ formal logic either. Formal logic and probability theory are indispensable when informal inferences need to be evaluated for rational acceptability—for validity or cogency. Since experimental inferences have conclusions that are more or less probable, the evaluative principles particularly appropriate to them should include the principles of probability. Bayes' theorem is a very important principle of this kind.

To understand how Bayes' theorem can be used in the evaluation and logical reconstruction of experimental inferences, we have to understand something about the principles of probability and how they can be applied to the task at hand. Studying probability theory can take you quickly into some serious mathematics, but the inferences I intend to describe here can be understood with only a minimal exposure to mathematical symbolism. I shall say just enough about the principles of probability to make an elementary use of Bayes' theorem understandable. You can understand me if you can recall the elementary parts of your high school algebra.

Formally speaking, the principles of probability comprise a remarkably simple mathematical system commonly known as the probability calculus. An important feature of this calculus (understood as a formal system) is that it can be interpreted in many different ways.[27] On one standard interpretation it applies to physical outcomes (changes in the world); on the one I shall use, it applies to statements or assertions. Applied this last way, the calculus concerns what are sometimes called epistemic or evidential probabilities.[28] As I shall understand them, these probabilities are degrees of certainty and evidential support.

The simplest probability statements of the kind in question are categorical in form; an example is "P(p) = a," which may be read "the probability of p equals a." The values assigned to these statements—for instance, the value represented here by "a"—are taken from the real numbers between 0 and 1 inclusive. 1 is the maximum value, indicating certain truth; 0 is the minimum value, indicating certain falsity. Since "$p \vee \sim p$" is certainly true and "$p \wedge \sim p$" is certainly false, $P(p \vee \sim p) = 1$ and $P(p \wedge \sim p) = 0$. The probability value of statements that are neither certainly true nor certainly false are represented by real numbers between 1 and 0; the value of statements closer to 1 are progressively more certain than those whose values approach zero and are progressively less certain than those closer to one. If we believe that a statement's degree of certainty is fairly close to 1, we might assign it a probability value of 0.9. If we think it is very uncertain, we might assign it a value of 0.2, which is tantamount to assigning its negation a value of 0.8.

---

[26] Another consideration, emphasized by van Fraassen, is that any inductive principle that yields conclusions incompatible with those obtained by the probability calculus (on the same evidence) possess a kind of incoherence. He supports this consideration by a so-called Dutch book argument, which I do not discuss in this book. See van Fraassen (1989), ch. 7. For further discussion of such an argument see Skyrms (1986), ch. 6.

[27] See Skyrms (1986).

[28] *Ibid,* p. 15.

These last remarks can be expanded to reassure readers not used to thinking of numerical degrees of certainty and support. Since certain truth is equivalent to a probability of 1 and certain falsity is equivalent to a probability of 0, a probability of 0.5 is equivalent to probabilistic indifference, where a statement is no more likely to be true than its negation. A probability of 0.75 is then intermediate between such indifference and certainty, so it amounts to "fairly probable" in everyday terms. Probabilities over 0.9 therefore count as "high." The fact that the probability assignments resulting from informal inferences to the best explanation are commonly thought to be no more precise than "slightly probable," "quite probable," and "highly probable" suggests that numerical assignments need not, in practice, be exact either. In most cases one can think of a numerical assignment as an approximation, representing a value in the neighborhood of what the number strictly represents.

Bayes' theorem provides a principle for calculating what are known as conditional probabilities. The formula "P(q/p) = a", as I shall interpret it, may be read "the probability of q on the assumption p" or, more simply, "the probability of q on p"; it expresses the degree to which q is evidentially supported on the assumption that p is true.[29] As in the case of ordinary deduction, the evidential support represented by conditional probability is hypothetical, because a false premise does not unconditionally support a conclusion that it entails. If Rover is a dog, the statement "Rover is a cat" entails "Something is a cat" but it does not succeed in showing that the latter is true. When we say that q is evidentially supported to the degree x by the hypothesis that p, we mean that the truth of that hypothesis would provide x degrees of support for q. (Other evidence that counts against q could, of course, undermine this degree of support if the latter is limited.)

Although, according to chapter one, evidence need not be propositional, a statement can always describe its nature or character. We can therefore use the term "P(h/e)" to denote the probability of h on the evidence e. If e entails h, the probability of h on e is maximal, as great as evidential support can be. Maximal support is represented by "1", the integer that also represents certain truth. As you would expect, maximal disconfirmation is represented by "0": if e entails ~h, P(h/e) = 0. Although conditional probability statements are ideally suited to express the degree to which a statement of evidence would, if true, support some hypothesis, they serve the more general purpose of expressing the degree to which one statement with a given probability value hypothetically supports another statement with a given value. Bayes' theorem, as I said, provides a general principle for ascertaining such a degree of support.

A simple form of Bayes' theorem can be stated as follows:

(SBT) If $P(e) \neq 0$, then $P(h/e) = P(h) \times P(e/h) / P(e)$.

This statement is much less complicated that it might initially appear. Evidence statements are almost always contingent statements, not certain falsities, so their probability is almost always positive. Thus the significant core of the theorem is the equality:

$$P(h/e) = \frac{P(h) \times P(e/h)}{P(e)}$$

[29] The formula "P(q/p)" is often glossed as "the probability of q given p," but the dangling participle in this locution is no clearer, in my opinion, than a dangler is in most other cases. I therefore prefer to avoid it here.

The left side of the equation can be taken to denote the probability of some hypothesis on the evidence e; the right side gives the formula by means of which this probability can be calculated, namely:

$$\frac{P(h) \times P(e/h)}{P(e)}$$

This last formula is not only very simple, but it encapsulates the commendable aspects of the reasoning in the hypothetico-deductive method and in inferences to the best explanation without including their defects. Unlike these other methods, it is sensitive to three things whose importance for hypothesis testing I have emphasized: the acceptability or antecedent probability of the hypothesis being tested (represented by "P(h)"), the antecedent probability of a predicted outcome (represented by "P(e)"), and the degree to which the hypothesis hypothetically supports that outcome (represented by "P(e/h)").

Suppose that the hypothesis h (into which we can incorporate pertinent background information) predicts e with a positive degree of certainty n. If we contemplate the fraction by which the degree to which e hypothetically supports h can be calculated—namely,

$$\frac{P(h) \times P(e/h)}{P(e)}$$

—we can see that the more improbable e is, that is, the lower the value of P(e), the larger the value of the fraction will be and, therefore, the greater the value of P(h/e) and the more strongly e will hypothetically support h. We can also see that the higher the credibility (or antecedent probability) of the hypothesis h, the larger the numerator of the fraction will be and therefore the larger that fraction will be. This means that (other things being equal) the greater the probability of the hypothesis, the more strongly it is supported by the occurrence of what it predicts. These facts accord beautifully with informal principles of a posteriori inference: implausible, ad hoc hypotheses that accord with observable data are not (generally speaking) strongly supported by that data; and the occurrence of antecedently unlikely data strongly supports hypotheses that predict it.

I mentioned earlier that if p entails q—that is, if it is certain that $p \supset q$—the conditional probability of q on p = 1. This principle holds for all p and q. As a result of this, if a hypothesis h entails e, the negation of e entails the negation of h—that is to say, ~e entails ~h. But if ~e entails ~h, P(~h/~e) =1. We therefore have a mathematically sound basis for the principle that the nonoccurrence of evidence predicted by a hypothesis effectively refutes that hypothesis, rendering its negation conditionally certain.

What about the problem of alternative hypotheses that I mentioned in the last section? When I discussed the hypothetico-deductive method, I observed that the occurrence of predicted data cannot accord significant support to a hypothesis by itself because data that accords with one hypothesis always accords with other hypotheses, thus adding no support to any hypothesis in particular. Does this observation undermine the usefulness of the version of Bayes' theorem we are considering? The answer is no. If, in computing the value of P(h/e), we assign a high antecedent probability to the hypothesis h, we single it out as a special hypothesis that (other things being equal) will receive significant support from data whose occurrence it allows us to predict. And if, on the other hand, we assign a low antecedent probability to e, the basis for our assignment can only be other hypotheses that we are tacitly

taking account of; we are in effect assuming that e has, on the average, a low probability given the totality of hypotheses bearing upon its value.

Ascertaining Prior Probabilities

The answer I gave in the last paragraph is bound to raise a more fundamental question: "How are we to ascertain the antecedent probabilities, the ones such as P(h), P(e/h), and P(e), that are needed to apply Bayes' theorem?" Actually, I have already given part of the appropriate answer. As far as P(e/h) is concerned, the appropriate value can often be obtained by ordinary deduction. If h is a compound formula containing not only a hypothesis but the background assumptions needed to support a prediction e, h can be expected to entail e, so that P(e/h) = 1. To make matters more perspicuous, we might identify these background assumptions explicitly, using a term such as "P(e/h ∧ a)" (meaning "the probability of e on h and a) rather than the "P(e/h)."[30] As in the simpler case of "P(e/h)," we can often calculate the value of a term such as "P(e/h ∧ a)" by ordinary deduction. If the conjunction of h and a entails e, P(e/h ∧ a) = 1; if that conjunction implies that e has a lesser value j, P(e/ h ∧ a) = j as well.

The antecedent probability of the categorical h's or e's required to apply Bayes' theorem can usually be calculated from pre-existing information concerning those assertions. Often the calculation for a hypothesis h can be made by a rule of conditioning (as it is called) applied to prior applications of Bayes' theorem. If, having used Bayes' theorem to calculate a value, say n, for the conditional probability of an assertion h on evidence e*—that is, for P(h/e*)—we may proceed to assign the value n to h itself—that is, to P(h)—if we learn that e* is in fact true. When we do this, we are updating the value of h. We needed an antecedent or "prior" probability for h in order to calculate its conditional value on the evidence of a predicted e*, but when we learn that "e*" is actually true we can give h a new "posterior" probability. This posterior probability can become a prior probability for further applications of Bayes' theorem; the designations "prior probability" and "posterior probability" are in fact applicable to an assertion only in relation to a Bayesian calculation and a verified prediction. Bayes' theorem is a powerful investigative tool because it can be applied again and again to a single hypothesis, updating its probability value as evidence accumulates. Thus, the value we assign to P(h) in an application of Bayes' theorem may often be computed by a prior application of that theorem to some prior, ascertained evidence e*.

As for the probability of an evidence statement e in an application of Bayes' theorem, we can often calculate this by estimating its value in relation to background hypotheses. If j and k are incompatible hypotheses one of which is bound to be true, we can estimate the value of P(e) by ascertaining its probability on both of these hypotheses and qualifying each conditional probability by factoring in the antecedent probability of each hypothesis. The rule to apply here, expressed symbolically, is "if j and k are jointly exhaustive, mutually incompatible hypotheses, then P(e) = P(j) × P(e/j) + P(k) × P(e/k)."[31] Here is a simple example of how the rule is applied. Suppose we have the following information about one member of a pair of dice: it is either fair or slightly biased in favor of heads, but much more likely to be fair than biased. We also know that the probability of getting heads if it is biased is

---

[30] Calculating the value of this more complicated probability requires a slightly more complicated Bayesian rule, specifically:

$$P(h/e\wedge a)=\frac{P(h/a)\times P(e/h\wedge a)}{P(e/a)}\text{, if } P(e/a)\neq 0.$$

[31] The antecedent here is "(j ∨ k) ∧ ~(j ∧ k)."

0.7 and that the probability of its being biased is 0.2. Consistency then requires us to assume that the probability of getting heads if it is fair is 0.5 and the probability that it is fair is 0.8. What, we want to know, is P(e), the probability of getting heads on a single throw of this die? The answer is $P(e) = P(f) \times P(e/f) + P(b) \times P(e/b) = 0.8 \times 0.5 + 0.2 \times 0.7 = 0.4 + 0.14 = 0.54$.

These strategies for computing values for the probabilities needed to apply Bayes' theorem have an obvious drawback from a philosophical point of view. They show us how to assign a probability value to a statement only if we already know other relevant probability values. These strategies do not therefore tell us how a basic probability value is rightly determined. Yet without basic probability values, we cannot assign a value for any probability other than a so-called likelihood—that is, a statement giving the conditional probability of an outcome on some hypothesis and background assumptions. (I have said that this kind of probability can often be determined by ordinary deduction.) How can we possibly ascertain basic probability values?

According to an influential school of statisticians known as subjective Bayesians, the basic probabilities needed for experimental inference can simply be assumed, because they do not have to be well founded or accurate in some sense. Experimental inference based on Bayes' theorem is, they say, self-correcting. If we begin with prior probabilities that are not extreme (close to zero or one) and continue to update our probability values by the rule of conditioning, the effect of our initial prior probabilities will become progressively smaller as we proceed: two people starting out with different prior probabilities and updating their probability values by successive conditioning involving the same evidential input will eventually agree on the probabilities they ascribe to relevant hypotheses. This claim, which can be demonstrated mathematically, shows that inferred probabilities can be acceptable without being based on objectively correct priors.[32]

The mathematical fact that people who update their prior probability functions by persistent conditioning on the same evidential data will eventually agree on the probability values they assign to resultant hypotheses does not really dispose of the philosophical problem at issue here. In actual cases in which the resultant agreement is approximated, there is a great deal of presupposed agreement on the admissible evidence, on the alternative hypotheses to be considered, and on such things as the probabilistic independence of occurrences pertinent to their calculations.[33] It is possible to seek experimental support for what is thus presupposed; but to obtain this support, further inferences of a probabilistic sort will have to be made, and these inferences will require further assumptions about prior probabilities and evidential data. If people with different priors disagree on any of these matters, the probabilities they eventually assign to the relevant hypotheses are not likely to be the same.

If we are to use Bayes' theorem as a basic rule of experimental inference, we must therefore find some way of justifying basic probability statements. If a statement is analytically true, it is of course certain and has a probability of 1; and if a statement p implies a statement q, the probability of q on p is also 1. Similarly, statements that are analytically false have a zero probability, and if p is inconsistent with q, the probability of q on p is also 0. In other cases, probability theory applied to statements cannot itself assign a value to any categorical statement.

It is important to realize that the limitation I have just mentioned also holds true for ordinary deductive logic: only logically true and logically false statements are given a definite value by logic itself. The value of contingent statements must be

[32] See Phillips (1973), Rosenkrantz (1981), Skyrms (1986), and Howson and Urbach (1989).
[33] I show this in the Appendix to Aune (1991).

determined empirically. As regards these statements, logic can tell us only what is true, or false, if something else is true, or false. Generally speaking, the point in knowing that Q is a deductively valid consequence of a premise P is that we should be inconsistent if we accept both P and ~Q. If these propositions concern matters of fact, the choice between them is not a logical one. If we accept P, we must not accept ~Q; if we accept ~Q, we must not accept P.

According to a subjectivist interpretation, the probability calculus places consistency conditions on statements expressing degrees of belief or confidence in propositions. As Frank Ramsey put it is his pioneering essay, "the laws of probability are laws of consistency, an extension to partial beliefs of formal logic, the logic of consistency."[34] This way of looking at the probability calculus can be illustrated by a fumbling attempt to apply Bayes' theorem. Since an antecedently unlikely occurrence strongly supports a hypothesis from which it is predictable, and since antecedently probable hypotheses are, other things being equal, more strongly supported by the predictions they warrant than antecedently improbable hypotheses, it might occur to a person beginning the study of probability that a favorite hypothesis h (to which he assigns a moderately high probability of 0.8) would be very strongly supported by a testable consequence e with a low antecedent probability of 0.4. A simple computation shows, however, that this arrangement of probabilities is inconsistent. If e is deducible from h (so that "h $\supset$ e" is certain), P(e/h) = 1. Given this value of the likelihood P(e/h), one can infer from Bayes' theorem that P(h/e) = P(h)/P(e), which = 2 in this case. But this is an impossible result, since no probability can be greater than 1. Reflection shows that if p entails q, the probability of q cannot be less than the probability of p—and this fact was not appreciated in the case I have described.

Basic Prior Probabilities

The fact that epistemic probabilities are constrained by analytic certainties is enough to show that a purely subjectivist interpretation does not accord with the approach I have taken here. It is not a matter of subjective belief that Q is analytically true or analytically false; it is also not a matter of subjective belief that a contingent R or S is known to be true. Contingent matters cannot be known to be true for certain, so they do not deserve a probability of 1; but anything that is known to be true in a looser sense deserves a very high probability--less than 1 but reasonably close to it. Since so called likelihoods—that is, assertions of the form "P(e/h$\wedge$a) = n"— can usually be given a probability value on the basis of deduction, the remaining probability statements that need some extra-logical justification are those assigning values to basic prior probabilities, those not inferred from other probabilities. The question is, "How can these basic probability statements possibly be justified?"

An answer to this question can be located by reflecting on a basic epistemic principle laid down by that critic of empiricism, R. M. Chisholm. The principle is:

CP: If S accepts h and if h is not disconfirmed by S's total evidence, then h is probable for S.[35]

The conception of probability involved in this principle is not the conception I have been concerned with in this chapter; it is an idiosyncratic conception that

[34] Ramsey (1931), p. 182.
[35] Chisholm (1989), p. 72.

Chisholm expresses by the adjective "internally probable." According to him, a proposition is internally probable for a person S just when S is more justified in believing the proposition than he or she is in believing its negation.[36]

If the conception of being justified that Chisholm employs is supposed to be closely related to truth, I would reject CP right away. There is surely no good reason to suppose that if anyone accepts something that is not disconfirmed by his or her total evidence, that proposition is apt to be true or even close to the truth. Ignorant and barbarous people, as Hume would say, believe all sorts of patently false things that are not disconfirmed by the evidence available to them.[37] Surely things so believed are not true more often than not. People pertinaciously obstinate in their delusions (another Humean turn of phrase) refuse to consider evidence ostensibly contrary to those delusions, so they are, in effect, insulated from anything that might disconfirm them. The mere fact that they are believed is hardly evidence in their favor.

Reasonable people who have a sincere interest in discovering the truth will not protect their illusions this way, so the fact that their beliefs are not disconfirmed by the evidence available to them is an epistemically much more significant fact. May we not suppose that such people have probably confronted ostensibly disconfirming evidence for their beliefs and ruled it out on rational grounds? May we not conclude that their surviving beliefs are apt to be true more often than not? I would say no; a general conclusion of this kind is excessively indiscriminate. A well-considered judgment on this matter must take into account contingent facts about these people and about the kind of beliefs they are apt to form. If, like the investigators Locke would have commended, they are in the habit of proportioning their beliefs to the evidence, they are unlikely to have beliefs for which they lack positive supporting evidence; their non-disconfirmed beliefs are probable only because they are rendered so by such evidence. Other people might be slightly less circumspect in forming beliefs, and some might not be circumspect at all. The latter may be willing to consider contrary evidence but they may readily form beliefs about domains for which there is little or no possibility of obtaining evidence: they may have elaborate theologies, mythologies, or fanciful histories that dominate their thoughts but have no testable consequences. The otherworldly beliefs that survive disconfirmation for these people may be almost invariably false; there is no reason to suppose that they are ever true.

Although Chisholm's unqualified principle CP is thoroughly unacceptable,[38] a qualified version strikes me as defensible. The first qualification concerns the word "acceptance." This is really not a good word for a properly qualified version of CP, because the notion of probability appropriate for such a principle is the degree of certainty notion I have been discussing, and the degree of probability appropriate to the "accepted" hypothesis must be minimal if the subject lacks positive evidence. For a properly qualified CP, the acceptance in question is best described as that of *weak acceptance on a trial basis*. The second qualification is closely related to the first one: the word "probable" must refer to a minimal degree of certainty. Since 0.5 represents probabilistic indifference, where the certainty of "P" is the same as the certainty of "~p," and 0.75 represents moderate certainty, the midpoint between indifference and certainty, minimal probability can be taken to be somewhere between 0.55 and 0.6. A probability in this range can be corrected by further testing involving Bayes' theorem. Its posterior value

---

[36] *Ibid.*, p. 87.

[37] Hume thought of some human beings as "extremely ignorant and stupid" and "ready to swallow even the grossest delusion." See Hume (1777), p. 120.

[38] For a criticism of Chisholm's defense of CP see Appendix 5.

will depend on the way it is supported empirically.

I have now explained how the probabilities needed to apply Bayes' theorem can justifiably be obtained. The priors needed for hypotheses are obtained by testing more primitive priors that were initially adopted as conjectures having predictive (Lycan would say "explanatory") potential; likelihoods are obtained principally by deduction from hypotheses and auxiliary assumptions; and priors needed for evidence statements are deduced from background information concerning the theoretical principles applicable in the case at hand. The results of predictions are ascertained, ultimately, by observation. In practice observations involve the application of theoretical knowledge, but this knowledge ultimately results from the observation of predicted results. When we update priors by applying the rule of conditioning, we are in effect assigning a probability of 1 to a verified observation statement. A lesser probability is, strictly speaking, appropriate, since empirical statements are always uncertain to some degree.[39] In practice this uncertainty is commonly disregarded, since precise probabilities are not generally required for most empirical investigation. We can always be more precise if we think we have to.

There are many issues pertinent to Bayesian reasoning that I have not touched on here; in fact, there are problems of varying seriousness that specialists in probability theory continue to debate.[40] One problem concerns old evidence. I observed that hypotheses are not strongly supported by the occurrence of events with a high prior probability, but sometimes there is a serious question about how old or familiar data is to be explained, and a hypothesis' ability to account for that data is assumed to count strongly in favor of it. Is such an assumption always objectionable? I would say no. If there is a serious question about how some familiar data is to be explained—if no available explanatory account is accepted as applying to it—then that data can reasonably be considered unlikely in relation to accepted principles, and statements reporting its occurrence can be assessed as such for explanatory purposes. Theoretically, the data, though actual and therefore probable, is surprising; and its surprise value is what is represented by the low probability value. A plausible hypothesis that can predict its occurrence is therefore increased in probability value; for purposes of the calculation its actuality is ignored. When the calculation is completed, the fact that the data is obtained may then be used to update the probability value of the hypothesis, making its posterior value a new prior. It seems to me that this procedure raises no significant problems. An investigator who supports a hypothesis by old data will of course want to test it further by means of new data, but the old data is more significant than it would be if it had to be assigned a probability close to one for purposes of the calculation.

Although there is much more that needs to be said about Bayesian inference,[41] what I have said is enough to show that it is far preferable to Inference to the Best Explanation, the alternative widely acceptable today. Apart from the actual defects that I have noted, IBE has no evident rationale in the first place: even lacking identifiable defects of the kind I mentioned, there is no evident reason why a rational person should be moved to adopt it. The same is not true of Bayesian inference. Bayes' theorem is a mathematically sound principle, inferable from axioms that, interpreted in the way I have suggested, are reasonably regarded as analytic.[42]

---

[39] See the helpful discussion in Rosenkrantz (1981), 3.6, 3-4.

[40] For discussion see Earman (1992) and the very perceptive review of Earman in Hájek and Skyrms (2000).

[41] See Talbot (2001), Howson and Urbach (1989), Skyrms (1986), Rosenkrantz (1981), and Phillips (1973).

[42] See Appendix 6.

This gives Bayesian inference a rationale not possessed by other principles of experimental inference. Of course, the results of a Bayesian inference are not analytic, for nonanalytic premises are required for the inference. But these nonanalytic premises are rationally defensible for the reasons I have given. Nothing analogous appears to be true for Inference to the Best Explanation, the only evident alternative for reasoning about matters that transcend the domain of the observable.

The BIV Hypothesis Again

If Bayes' theorem provides a valid form for experimental inference and if, in addition, we may justifiably accept (under the conditions I mentioned) observational premises, prior probability assignments, and delimited sets of alternative hypotheses, the way is then open for confirming the facts about observers that support or refute what they claim to have observed or remembered. More than this is actually possible: Means are now available for dismissing Putnam's BIV hypothesis, and doing so by straightforward empirical means. Since this latter task is more fundamental than that of supporting pertinent facts about observers or those purporting to remember something, discussing the strategy for dismissing the BIV hypothesis is a fitting way of ending this chapter.

To show that Putnam's hypothesis deserves to be rejected on empirical grounds, we need only show that an alternative, realist hypothesis—one formulating our best estimate about the actual nature of ourselves and our world—is better supported, empirically, than the hypothesis about BIVs and their relation to the computer that orchestrates their delusional experiences. This would seem to be a very easy thing to so, because the BIV hypothesis represents a mere conceptual possibility, one that, like Descartes' hypothesis of an evil genius, has absolutely no evidence in its favor. If we can show that the realist hypothesis we actually accept has *any* evidence in its favor, we would evidently accomplish this task in short order. Unfortunately, the issue is not quite this simple. If we take the BIV hypothesis seriously as even a possibility, we will be thinking of our available evidence as consisting of facts about the subjective experience of the relevant subjects. So the question for us to consider is whether, by reference to that kind of experience, we can show that our realist view is better supported than Putnam's BIV hypothesis.

To support the view that we actually accept—it sounds odd to call it a hypothesis, but I shall do so for the sake of argument—it is convenient to begin with a simpler, more specific hypothesis, ST, one that entails some of the consequences of the accepted alternative for my own current physical situation in the world. ST asserts something very specific--namely, that I am (really) sitting at my desk in my study looking into my word processor and that there is behind me a Bertoia chair covered with red upholstery. As I implied at the beginning of this section, Bayes theorem permits us to show that this simple hypothesis can readily be confirmed by reference to my subjective experience. The procedure is straightforward.

Given ST and some related assumptions concerning the nature of what I take sitting, looking, seeing chairs, and acting to be, I can justifiably predict that if I will to turn around 180° and subsequently have the experience of doing so, I will in fact turn around this way and subsequently see, and so have the experience of seeing, the red Bertoia chair. I so will and I have the predicted experiences. The likelihood here, the probability of having the indicated experiences on the supposition that the hypothesis and auxiliary assumptions are true, can reasonably be set very high: the hypothesis and assumptions can be elaborated in a way that warrants this prediction

with a high degree of certainty. The antecedent probability[43] of having the experience of seeing a red Bertoia chair if I merely will to turn around can, by contrast, be reasonably considered quite low, since willing to turn my head is rarely followed by such a visual experience: it is so followed only in cases when I believe I am in my study looking into my word-processor. What then about the prior probability of my hypothesis, ST? If I am not to beg the question against BIV, I cannot make use of the evidence I would normally advance if I if were asked to defend my conviction that I am at my desk in my study facing my word processor. That evidence is based on defeasible presumptions that are now in question. Without that evidence, I cannot support a high prior probability for ST; I can only give it a moderately low prior probability in accordance with the strategy I described when I mentioned Chisholm's principle CP. But even with a moderately low prior for ST, a high value for P(e/ST) and a low value for P(e) yields a high value for P(ST/e) and, in view of the truth of e, a high posterior value for P(ST).

Of course, since the prior for ST was assumed on a trial basis, further tests are required, possibly with slightly different priors chosen for ST. But if the predications made on the basis of these different priors are satisfied to a comparable degree (as they would be, since the predictions would be the same), the posterior values for ST would remain high and justify a high value for that hypothesis. Since BIV is incompatible with ST, a high posterior value for ST requires a low value for BIV. Thus, BIV is effectively ruled out in favor of ST.

A supporter of BIV might object to this reasoning, saying that the antecedent probability of BIV deserves to be set as high (for purposes of the argument) as ST and that it would be confirmed just as strongly as ST since it will warrant the same predictions as ST. This objection fails, however. It is true that on a generous interpretation of what BIV asserts—one specifying appropriate links between the intentions and know-how of the scientific maniacs, the computer program, and the experiences of the disembodied brains—BIV will predict everything that ST predicts. But BIV makes many untestable claims that ST (or the realist hypothesis that ST represents) does not make, and this surplus content requires it to have a lower prior probability than ST—far lower, in view of its extravagantly rich sci-fi content.

Why does the surplus, untestable content of BIV require it to have a lower prior probability than ST? The reason is this. If BIV has every testable consequence that ST has but not vice versa, then BIV can be divided into two parts, A and B, one of which, A, represents the untestable content of BIV and the other, B, represents the part that is empirically equivalent to ST. Now A is probabilistically independent of B; P(A on B) = P(A). (If A and B are independent, the probability of either is unaffected by the truth of the other.) We knows that A and B are independent because B represents the part of BIV that does not have testable consequences, and A represents the rest of it, which has the testable consequences of the fully testable RT. But if A and B are independent in this way, P(A ∧ B) is less then both P(A) and P(B) if A and B are neither analytically true nor analytically false (as they are in the BIV case). This is a consequence of the theorem for the probability of such conjunctions: If A and B are independent, P(A ∧ B) = P(A) × P(B).[44] Thus, owing to its untestable part B, the prior probability of BIV is less than the prior probability of ST. Given Bayes' theorem, we can therefore conclude that ST is more strongly confirmed by its supporting evidence than BIV is by its supporting evidence. By hypothesis, the supporting evidence in both cases is the same.

---

[43] This probability is actually conditional on the assumptions about sitting, looking, seeing chairs, and acting that I have mentioned. I am here using the more complicated version of Bayes' theorem set forth in footnote 27.

[44] When two fractions are multiplied together, the product is less than either fraction. 1/2 × 1/2 = 1/4.

It is vital to realize that simplicity is not the only consideration that matters here. Equally important is the fact of empirical equivalence, the fact that the rival theories have the same testable consequences. Someone might argue that the ST hypothesis may be simpler, in the indicated way, than the BIV hypothesis, but that other rivals to ST may be even simpler. One such rival is a form of phenomenalism, the doctrine that only experiences are fundamentally real and that words such as "person" and "physical body" refer to nothing other than aggregates of, or constructs wholly reducible to, experiences.[45] There is no doubt that a theory of this kind is ontologically simpler—simpler in the sense of postulating fewer irreducible objects—but no such theory has ever been worked out in a satisfactory way,[46] and the best examples of the kind have been acknowledged to be inadequate by their authors.[47] Apart from this, the testable consequences of such a theory are by no means the same as those of ST: the latter, even the limited representative of the full-blown realist hypothesis that we actually accept, concerns the experiences of this or that person, and a person is a subject of experiences, not an aggregate of them.

I do not wish to plunge into the swamp of suggestions and replies that were once pursued when "Our Knowledge of the External World" was the leading topic on a philosopher's agenda. My argument against the BIV hypothesis was prompted by an imagined challenge--that of showing that skeptical hypotheses that are possible alternatives to the realist views we normally accept are not uncritically brushed aside but deserve to be rejected for identifiable reasons. I have tried to put my finger on some of those reasons. Much more could be said; the ST hypothesis as I described it was not the full-blown realist alternative to the BIV one: it was merely a limited hypothesis about me in my study, not a full-scale hypothesis about the nature of the world I inhabit and the creatures that share it with me. To move from ST to the alternative I have inadequately sketched, I should defend more of the presumptions we commonly make in thinking about the world. One is that the animated bodies I call people do not just behave intelligently and smile or frown when they are stroked or poked, but also think and feel much as I do. This presumption, also high on the philosophical agenda at another time, can be defended as well by a Bayesean strategy, but I have no interest in pursuing it here.

Historically, philosophers presented with skeptical hypotheses of the kind I have been considering have tried to refute them—to rule them out—by some kind of a priori strategy. They are cognitively meaningless; they fall short of the requirements for objective reference; they presuppose a nonsensical private language; they violate the sound requirements for the acceptable interpretation of meaningful discourse; and so forth. But according to the standards of a reasonable empiricism, these hypotheses are clearly meaningful. We know exactly what they mean; they would not be problematic for us if we could not understand them. They are, of course, far fetched, but that is not enough to show that they are false. What entitles us to reject them is that they are not nearly as well supported by available evidence as the hypotheses we accept. This evidence supporting accepted hypotheses is not perfect; it does not render them acceptable beyond the shadow of a theoretical doubt. Yet it is sufficient for knowledge in the sense that we commonly employ. If

---

[45] Fumerton (2005) suggests that Berkeley's hypothesis about God would be a simpler hypothesis, but phenomenalism would be even simpler and would not involve the enormous untestable surplus of Berkeley's hypothesis.

[46] I outline my objections to phenomenalism in Aune (1991), chapter four, section 8.

[47] Bertrand Russell gave up on phenomenalism, which he had developed in *Our Knowledge of the External World* (1914), as early as 1927, when he published his *Analysis of Matter*; Carnap abandoned the phenomenalist program of his *Logische Aufbau der Welt* (written 1922-25 but published in 1928) as early as 1936, when he published "Testability and Meaning." For discussion, see Freidman (1999), ch. 5.

we recognize only perfect knowledge, we will have to cope with a form of skepticism.[48] But there is no actual need to proceed this way.

When I hear philosophers seriously endorsing a skeptical view of human knowledge, I think of a photograph I recently saw of the Martian landscape, one of a series taken by the Rover vehicles in 2004 and transmitted back to earth. The photo looked a lot like a photo of the Mojave Desert. The color of the landscape was different, but it seemed very similar nevertheless. When the photo comes to mind, I am struck by the extraordinary achievement it represents. The engineers who created the Rover vehicles and the rockets that carried them to Mars had to possess an enormous amount of detailed knowledge about a bewildering variety of phenomena, and those who carried out the missions that produced the photos had to be right in more calculations than I could possibly enumerate. Yes, there are many things we do not know about our universe and ourselves, but there is an astonishing amount that we do know very well. It is sometimes hard to believe that the creatures who were hunting with arrows and spears ten centuries ago can now send robots to distant planets and then later leisurely view in their home television screens the photos sent back to them, doing so as they sip a cup of coffee or drink a glass of wine.

Concluding Remarks

The empiricist epistemology I have defended in this book is partly classical, partly reformed. Its basic structure is largely classical, recalling the empiricism of David Hume. Corresponding to Hume's division of the objects of human inquiry into "relations of ideas" and "matters of fact and existence," I have defended a distinction between analytically true and synthetically true statements. My distinction is not categorical, however; it does not place every true statement into one of two disjoint classes. Natural dialects or even idiolects are not sufficiently determinate to allow such a distinction, but careful speakers seriously concerned about the precision and truth-value of what they say can make their meaning sufficiently determinate to place the important things they want to say into one or the other of these classes.[49] Assertions that are analytically true are either logically true or A-true in Carnap's sense. This conception of analyticity rests on a distinction between the theorems of some assumed system of formal logic and assertions warranted by stipulative explications or determinate verbal conventions. The rationalist's supposed example of a synthetic a priori truth, ""Nothing can be both determinately green and determinately yellow all over," is actually an assertion of this last kind; it is in fact inferable, as I show, from a verbal convention about what counts as a single determinate color. Spelling out all the distinctions and qualifications needed to extend an analytic/synthetic distinction to thoughts as well as statements requires a fairly elaborate exposition; what I have just said is merely a skeleton summary of points I defended in chapters two and three.

Another classical element of my empiricism is the list of basic sources by which we can ascertain matters of fact and existence. Hume referred to these sources as observation, memory, and experimental inference. Each source is beset by distinctive problems, but each is unquestionably a source of genuine knowledge, at least when measured by ordinary, imperfect standards. A key component in all three sources is human experience, which is fundamentally directed to a world of things and persons. It is by hearing my fiddle that I know it to be in or out of tune; it is by tasting my wine that I know it to be Zinfandel, and it is by looking at a clock that I know what the time is supposed to be. The knowledge of the world that we get

[48] As Fogelin (1994) in effect argues.

[49] Mathematical assertions are possible exceptions; see my chapter two, p. 74f and footnote 2.

from the three sources is not certain, but it can be improved by other data. Initial probabilities are transformed into progressively more certain convictions. Because empirical confirmation is a dynamic process, taking place in many different directions, the totality of what we know at any particular time is inevitably somewhat disorganized; we can only strive for greater unity. As a consequence of these facts, popular depictions of the structure of what we know—foundationalism, coherentism, or even Quine's holism—are all inaccurate in some significant way.

As I have emphasized, empirical knowledge as we commonly understand it is not the kind of knowledge that Descartes was after. It falls far short of rational certainty. When knowledge is understood this last way, far-fetched possibilities such as Descartes' evil demon or Putnam's brains in a vat assume an epistemic importance they do not really deserve. We have no way of *proving* that these possibilities are not actual—that the stories they involve are in fact false—but we can show that they are far-fetched and that the contrary views we actually accept are much more likely to be true. And this is enough for a reasonable empiricism, one appropriate for a philosopher who aspires to be tough-minded but epistemically up-to-date. Classical empiricists often emulated Descartes in their quest for certainty, and skeptical scenarios have therefore persistently threatened them. We can avoid this outcome by a more realistic estimate of the kind of certainty we can hope to achieve.

Two assumptions once thought distinctive of a responsible empiricism must be firmly set aside. One is the assumption that our empirical knowledge or well-founded opinion must rest on a foundation of subjective experience. Not only does our empirical knowledge fail to rest on anything that deserves to be called a foundation, but the nature of our subjective experience is also, as I noted, quite questionable, generating on-going controversy among philosophers and even empirical scientists. The other objectionable assumption is that inherently unobservable objects are unknowable and cannot meaningfully be described or referred to. The classical view of meaning and legitimate reference on which this assumption is based is simply untenable. Meaningful words or ideas need not, as Hume said, be "traced back to original impressions";[50] they arise from inbuilt neural mechanisms and goal-directed, usually cooperative behavior,[51] neither of which is understandable by reference to immediate impressions.

Because the use of Bayes' theorem provides a rationally acceptable means of confirming hypotheses about objects and process that fall outside the domain of the observable, knowledge as we commonly understand it need not be restricted to observable phenomena. Saying this does not commit me to the view that acceptable scientific theories must always be interpreted "realistically," as describing actually existing, mind-independent objects. Acceptable theories can do different things: some can provide mere models or vehicles of prediction, which do not purport to describe anything observable. The most that we can generally require of an empirical theory is that it be, in van Fraassen's words, "observationally adequate."[52] But a general requirement of this kind does not preclude a realistic interpretation of some theories. That would be going too far. Empirical reality can contain many things too remote or too small to be perceived by us, and it can contain numerous things that we could never observe for reasons that we cannot even anticipate. A reasonable empiricism prescribes experience as our ultimate basis for empirical knowledge, but it does not limit our knowledge to things that we can observe or otherwise experience.

---

[50] Hume, *Enquiry*, section II.

[51] The support for this is now given by scientists rather than philosophers; for a general discussion see e.g. Pinker (1994).

[52] van Fraassen (1980).

APPENDICES

Appendix 1

Armstrong's New Hypothesis about Universals

A few years ago a visiting lecturer at my university publicly rebuked me for my lack of belief in universals. "Isn't it obvious that there is a such a thing as love?" he asked. Although I would now say that I do accept a theory of universals—namely, a distributive view of F-universals—I think the speaker's question was nevertheless inane. Some philosophers might have reservations about the example the speaker used, but most seem to share his belief in some sort of irreducible universals, though they are usually unable to say much about them. The existence of such things is generally assumed without question these days; those who seriously discuss it tend to be treated with incredulity or contempt. In a recent paper D.M. Armstrong suggests changes in the theory that I criticize in chapter four; the changes show that one serious defender of an A-theory of universals is fully aware of the problems such a theory faces.[1] They also underline the inanity of the speaker's question about a "thing" called love.

People unfamiliar with historical quarrels about universals might agree with the speaker that there is such a thing as love. But they will not suppose that love is some kind of entity. When they say yes to his question they will be thinking that people do fall in love and commonly love their mothers, their children, or their pets. The only objects or entities they will be thinking of are persons, and they will have things to say about how persons feel and behave in regard to the beings they love. But behaving and feeling are not objects. Apart from persons and their feelings and behavior when they love, nothing else comes to mind when ordinary people think about love. There is no further "thing" to think of.

Armstrong, in his new paper, expresses no personal doubts about the existence of universals, but he acknowledges that the theory he developed earlier faces problems, at least one of which is quite serious. The serious problem concerns the relation between particulars and universals, a subject I discuss in chapter four. But Armstrong believes this problem may have a solution, and he devotes his paper to a discussion of it. The basic idea of the solution is that particulars and the universals they are said to "instantiate" are unities that intersect; the intersection is a "point of partial identity."[2] Armstrong's "rough but perhaps helpful model" for the intersection is a cross that has been cut out of a single piece of wood. The intersection of the vertical and horizontal portions of the cross models the intersection between a universal and a particular that instantiates it.[3]

Armstrong's new theory raises many issues, more than I can discuss in this appendix. But it has one virtue, a limited one in my view, which I want to comment on here. Most theories of immanent universals leave the nature of universals significantly mysterious. Suppose $scarlet_{23}$ is a genuine universal. (As I observed in chapter four, Armstrong takes universals to be absolutely determinate realities.) Although we might be quite familiar with $scarlet_{23}$ particulars (many balloons may be colored that way), we might wonder what the universal $scarlet_{23}$ is like. Does it have the same color as the particulars that instantiate it? Armstrong will now say yes. The particular is what it is because of the universals with which it intersects. Since

Copyright 2008 by Bruce Aune

[1] Armstrong (2004B).
[2] Armstrong attributes the solution to Donald Baxter, who worked it out in Baxter (2001).
[3] Ibid, pp. 140f.

the intersection is a region common to the particular and the universal, the qualities of the particular in that region are exactly the same as the qualities of the universal there: if one is colored there—if it is $scarlet_{23}$—so is the other. This removes some of the mystery about the nature of universals as Armstrong understands them, but it raises difficulties of its own. If love is a universal, is part of it identical to part of a lover? But what would that part be like? Would it contain a feeling or a smile? Could anyone confidently answer these questions? A deeper, more significant question is "What reason could there possibly be to postulate such universals?" You can't expect to explain why a marble is red by postulating a higher-order object that is also red. If a particular's color needs a metaphysical explanation, the same is true of a universal's color—and so on, as I show in chapter four, up an endless ontological ladder.

The new theory has other problems. If a $scarlett_{23}$ particular is destroyed, the universal of which it is a part would be partly destroyed and therefore changed. Would this change simultaneously alter all the other particulars that instantiate this universal? They certainly would not be what they are in respect of color by virtue of their intersection with the very same unaltered universal. Armstrong is admirably tentative in his paper; he does not insist that his new theory will succeed. It may be better than I believe it is, but if universals are what it says they are, the point in postulating them is hard to understand. It is far from *obvious* that any theory of this kind (any A-theory) is even slightly promising. My critic's appeal to the obvious truth of a theory of universals is off the wall.

It should go without saying, but it is probably advisable for me to say it anyway, that you can speak of properties and relations without accepting some theory of universals. Speaking of such things can merely be a convenient *façon de parler*. If I wish to speak generally about how this or that particular thing is or looks, or how it and other things are related to further things, I can speak of their properties and relations. And, like Heil (2004), I can discuss the question whether a thing's qualities (what it is like qualitatively) are or are not powers (a disposition to affect other things in certain ways) without supposing, or being committed to accepting the idea, that qualities and powers are some kind of abstract object. The fact that anyone who discusses philosophical issues is almost compelled to use the language of properties and relations does not imply he or she should acknowledge that this language applies to some special domain of objects not dreamed of by people who have never studied philosophy.

Appendix 2

BOGHOSSIAN AND FIELD ON BASIC LOGICAL PRINCIPLES

My strategy for justifying basic logical truths and rules of inference by reference to semantical rules would be rejected by some of today's empiricists. Recent articles by Paul Boghossian[4] and Hartry Field[5] provide good examples of this disagreement. Although I am convinced that my argument in the text provides adequate support for the conclusions I draw there, some readers will want to know how I would respond to such critics as Boghossian and Field. I outline my replies here. They introduce material that adds perspective to what I say in the text.

[4]See Boghossian (2000).
[5]Field (2005.)

Boghossian argues that the basic truths and rules of classical logic cannot be justified without circularity. He does not conclude that these truths and rules must therefore be regarded as unjustifiable by anyone who is not a determined and level-headed skeptic; instead, he contends that a modest degree of circularity—one sufficient to justify the basic truths and laws by inference--must be regarded as acceptable by those who are not determined skeptics. Obviously, this is contrary to what I argue in chapter three.

In supporting his position Boghosian argues that if our sole primitive rule of inference for our system of propositional logic were modus ponens (MP), any attempt to justify this rule by means of truth tables, as I do in the text, would be circular. He tries to show this by giving a specimen truth-table analysis that employs an MP rule in an attempted justification. But there are several serious defects with his argument. First, if modus ponens is, loosely speaking, the sole inference rule of a system we are using, we are not restricted to this rule when we reason about that system. So if we wanted to show by a truth-table test that MP is truth preserving in our system, we could legitimately do so by using alternative rules such as disjunction elimination and double negation. Thus, "p ⊃q" is by definition true just when ~p ∨ q; "p" is true just when ~ ~p; so "q" is true when both premises are true.

In addition to this initial slip Boghossian errs in seeking a justification for an all-encompassing MP rule anyway. As I explained in the text, a basic rule of logic can be successfully justified only by reference to a class of permissible substituends for its schematic letters.[6] To know that a rule is truly truth-preserving or that a formula is universally valid in a classical system, we have to know that the permissible substituends for its schematic letters will not cause trouble, as they might well do if they involve vagueness, non-designating singular terms, and predicates such as "is true" and "is false." Classical systems set up in the right way are provably *consistent* (contradictions are not derivable), *sound* (whatever is provable in the system is logically true), and *complete* (all logical truths expressible in the system can be proved in it).[7] The proofs that classical systems have these features were important discoveries and are not circular: the logic used for the proofs is not the same as the logic of the object language if for no other reason than that its variables do not have the same substituends. The substituends are not extraneous to a logic principle; they are part of its content. The logic we use in reasoning about a logical system is logically distinct from it, no matter how much the two might look the same when they are considered in abstraction from their application.

When we use a truth table or a truth-value analysis to show that a formula of a classical system is a theorem or that an argument form is valid, we normally proceed informally, using vernacular terms such as "if" and "and." These terms correspond only approximately to logical symbols such as "⊃" and "∧." "If" is particularly loose as ordinarily used; and "and" can not only appear between expressions of categories not allowable for "∧"—such as nouns, verbs, adverbs—but it is sometimes used in the sense of "and then," which does not satisfy the principle of commutation for conjunction. In spite of these peculiarities of vernacular counterparts to logical symbols, elementary inferences carried out in everyday language, such as "If A and B, then A" (where "A," "B," and "C" are specific sentences), may reasonably be considered analytically truth-preserving—just like the inference "Tom knows that snow is white, so it is true that snow is white." Being truth-preserving is not the same as being *formally* valid. Formal validity makes sense only in relation to a class of for-

---

[6] The initial explanation occurs in chapter 2, pp. 32-36. See especially the quotation from Gödel on pp. 34f.

[7] See Church (1956) for proofs of these contentions. He discusses the concept of soundness on p. 55 and consistency and completeness on pp. 108 and 109.

mulas that share a form. When we reason informally, we have no such class in mind. Nevertheless, we may be fully justified in holding that a particular inference (not an inference schema) involving a particular vernacular connective will not take us from true premises to a false conclusion. Our conviction on this matter will not be rationally enhanced by relating the single vernacular inference to a wider class of inferences. It may be sufficient unto itself. On this see footnote 39 of chapter three and the text above it.

My view here actually has some affinity with the position of Hartry Field outlines in his article. Field claims that we are "default-entitled" to our logic and methodology; in the absence of considerations specifically raised against it, he says, we are entitled to use this logic, whether deductive or inductive, without having any argument for it.[8] I agree that we possess this default-entitlement; it is of a piece of our normal entitlement to use the methods and principles that I spoke of in chapter one when I discussed knowing in the weak or "loose and popular" sense. But the authority and ultimate reliability of these methods and principles is a proper subject of investigation in epistemology. This investigation need not be seen as concerned with our justification for accepting these methods and principles but with how and to what extent their reliability and truth can be ascertained. As I show in chapter six, some of the so-called inductive methods commonly accepted even by philosophers are in fact defective and require significant qualification.

Field, in spite of emphasizing that we are default-entitled to standard logic principles, goes on to suggest in the same essay that debates about the acceptability of various logical principles can be settled by holistic considerations. "The consequences of changing logical opinions," he says, citing excluded middle as an example, "can be far-reaching, and we need to look at quite diverse consequences of the change and decide whether the benefits to our all-over world-view would outweigh the costs" (p. 87). But if excluded middle is formulated in the usual symbolic way—schematically as "p $\vee \sim$ p"—and admissible substituends for the schematic letters are restricted, as they have to be, to statements that satisfy the principle of bivalence (they are either true or false but not both), then the principle cannot be falsified. A sentence with "$\vee$" as main connective is stipulated to be true just when at least one of its ingredient disjuncts is true, and if that disjunct is true (or false) bivalence guarantees that its negation is false (or true). Bivalence is not just a meta-principle that the formulas of a system are *conjectured* to have; it is a condition that a sentence must meet to be a proper substituend for the excluded-middle schema.

If we change our criteria for being a proper substituend and include formulas not satisfying bivalence, we will of course, in a sense, be rejecting excluded middle. But the principle as we understood it before we changed our criteria will remain sound—just as syllogisms in the Barbara form remain valid even though we no longer use the syllogism as a component (or the whole) of our logical theory. (We no longer think of "All sheep are animals" as having the form of a syllogistic A-statement, "sAp," or of a Boolean identity, "$\alpha \cap -\beta = 0$.) Moving beyond excluded middle in this way is not motivated by our all-over worldview, as Field suggests; it results from the decision to include substuends that, owing to such things as vagueness, may lack a determinate truth-value. Some experts on quantum mechanics have argued that a logic compatible with that theory should lack a distribution law,[9] but even if they are right (other experts deny this), a special logic for quantum mechanics is compatible with a classical logic for other purposes—just as a logic with an

[8]Ibid, p. 86.

[9] See Wilce, Alexander, "Quantum Logic and Probability Theory", *The Stanford Encyclopedia of Philosophy* (Spring 2006 Edition), Edward N. Zalta (ed.), URL = http://plato.stanford.edu/archives/spr2006/entries/qt-quantlog/>..

"and-then" connective,[10] which lacks a commutative principle, is compatible with a classical system featuring a connective for ordinary conjunction, which permits commutation.

Appendix 3

STIPULATION, PROPER DEFINITIONS, AND TRUTH

Paul Horwich (2000) vigorously attacked the idea, which I defend in the text, that a priori truths can be created by stipulation. His attack was focused, however, on the sort of stipulation used by Paul Boghossian in expounding his version of what he called "the analytic theory of the a priori." Although I defend analyticity in this book, I am not sympathetic with Boghossian's version and I am equally unsympathetic with the kind of stipulation he employs. Thus, I have no objection to Horwich's criticism of stipulation so understood. The kind of stipulation I introduce in connection with Carnap's notion of an explication is quite different. In this appendix I formulate my own objections to the kind of stipulation that Boghossian used (they are different from the ones Horwich raises), and then I make some remarks about the notion of a "proper" definition. My remarks about the latter notion have implications for the sort of stipulation, or partial explication, that Carnap introduced in connection with his A-postulates. This kind of stipulation is not vulnerable to Horwich's perceptive criticism.

The sort of stipulation that Boghossian introduced and Horwich criticized has the following form:

> The word ϕ is to have whatever meaning will make true a certain conjunction of postulates, $\Pi(\phi)$, containing that word.

As I explained in the text, the sort of stipulation I favor is intended to specify, either wholly or incompletely, how some expression (some word or phrase) is to be understood in a particular context or discussion. But a person making a stipulation of the kind displayed above need have no determinate idea of the meaning he/she intends the target word or formula to possess, and the stipulation he or she offers cannot be expected to bring a determinate meaning to the attention of a potential hearer. The reason for this is that any number of different meanings for φ could suffice to make some postulates containing φ true, and to assert that it may have any one of these meanings is not to specify any particular meaning for it.

Suppose I wish to stipulate the meaning the word 'lunatic' will have in a discourse that I will provide, and suppose that I offer just one postulate, namely:

> d. $\forall x$(x is a lunatic $\supset$ x is mentally deranged and x's derangement has a lunar cause).

It is obvious that the formula (P) could be made true by countless meanings for 'lunatic.' Here are just two: 'x is a mentally deranged horse and x's derangement has a lunar cause'; 'x is a mentally deranged philosophy professor and x's derangement has a lunar cause.' Obviously, the number of expressions that could be substituted for 'horse' and 'philosophy professor' here is endless.

---

[10] A logic of this kind was once used in von Wright (1971).

The objection I have just made could be avoided if the postulates $\Pi(\varphi)$ were required to formally entail a biconditional sufficient to provide an explicit definition for the expression φ. But if $\Pi(\varphi)$ satisfies this condition, $\Pi(\varphi)$ is either equivalent to that biconditional or contains additional terms or clauses that are not pertinent to the meaning-fixing content of $\Pi(\varphi)$. In either case a biconditional is available that can provide a stipulative definition of the usual kind for the expression φ. A Boghossian style stipulation is neither needed nor wanted.

There are many things I cannot possibly stipulate. I cannot stipulate what you or someone else shall mean by an expression; I cannot now stipulate what I once meant in using an expression E (though I can certainly disclose it); and I cannot stipulate that something false or inconsistent is true. Thus, there are many restrictions on a successful stipulation. I have perhaps touched on one of the most basic restrictions already: the formula I use in making a stipulation must have an appropriate structure. As I have indicated, one such structure is that of a defining biconditional. An instance may be fashioned from the fanciful example I used earlier:

> ∀x(x is a lunatic ≡ x is mentally deranged and x's derangement has a lunar cause).

In using a biconditional of this kind one may stipulate that the predicate appearing on the left-hand side of the formula will be used to apply to a thing x, should such a thing exist, just when x is mentally deranged and when there is a lunar cause of x's derangement. Such a stipulation will plausibly generate an a priori truth because if anything does satisfy the specified condition—if, that is, it is mentally deranged and there is a lunar cause of the derangement—then and only then will the introduced predicate apply to it: it will then be, in the intended sense, a lunatic. The relevant biconditional is therefore true.

As I mentioned when I discussed Carnap's reduction sentences, one may wish to specify an expression's meaning only incompletely. In such a case one's stipulation is partial rather than exhaustive or complete, and the formula introducing one's stipulation may have the form of a mere conditional. An example may be given using the formula (P) above:

> ∀x(x is a lunatic ⊃ x is mentally deranged and x's derangement has a lunar cause).

Here one specifies a condition that something must satisfy if the introduced predicate is to apply to it—if it is to be, as this case requires, a lunatic in the intended sense. The formula gives no rule for identifying an instance to which the predicate does apply, so the formula is not as helpful as it could be. But one might be interested merely in clarifying a particular aspect of how one will be using a common expression, leaving other aspects to the judgment of one's hearers or readers. In this sort of case a partial stipulation may be adequate to one's needs.

In clarifying certain aspects of a word that, as commonly used, is significantly vague, one might introduce more than one partial stipulation. Take a word like "thin." Obviously, there are many borderline cases to which this word neither clearly applies nor clearly fails to apply. But one could lay down conditions that permit others to identify certain clear positive cases and certain clear negative cases in relation to one's intended use of the predicate:

> ∀x(x is gaunt or scrawny ⊃ x is thin).
> ∀x(x is thin ⊃ ~x is obese).

To generate statements that deserve to be considered analytically true of a person's discourse in some specified context, attempted stipulations must satisfy further conditions. One is that the proffered formulas must be consistent—both internally and in relation to other rules and stipulations to which the speaker is committed. Consider the following formulas, which might have been used in an attempt to provide stipulative introduction and elimination rules for the expression "tonk":

(1) $p \supset p \text{ tonk } q$.

(2) $p \text{ tonk } q \supset q$.[11]

Since these formulas are supposed to hold for all permissible substituends for the schematic letters 'p' and 'q', an explicit contradiction is easily deducible from them.[12] They do not therefore provide admissible meaning rules or stipulations.

This last result would hold true even if a contradiction were not derivable from the two rules. Another condition that a proper definition or stipulation S should satisfy, whether it is complete or only partial, is that it not be creative in the sense of implying something T that (a) is expressible in the language of the system before S is introduced and (b) that is not implied by the logic and semantical rules of the system independently of S. The aim of a definition or meaning specification for predicates or formulas is merely to provide truth- or satisfaction-conditions for those particular symbols; they should not introduce new truths about the subject matter in question.[13] But rules (1) and (2) permit the deduction of an additional formula "$P \supset Q$," which can be understood as expressing a contingent truth about the relevant subject matter. A proper definition, stipulation, or meaning specification should not do this. It violates a criterion of non-creativity that such formulas should satisfy.

There are further conditions that definitions and stipulations should also satisfy, but there is no need to discuss them here. The best, most complete discussion of the subject known to me is contained in a chapter to which I have already referred, Suppes' "Theory of Definition."[14] My aim here is merely to indicate that stipulations can create verbal truths only when they are *properly formed* and *properly introduced.*

---

[11] These formulas are mentioned by Horwich (p. 155); they were first used by A. N. Prior in Prior (1960).

[12] Taking different substituends for the schematic letters in these formulas, we may infer both $P \supset \sim P$ (which is equivalent to $\sim P$) and $\sim P \supset P$ (which is equivalent to P).

[13] See Suppes (1967), p. 154.

[14] Ibid, chapter 8.

Appendix 4

ON "WHAT IS SAID" AND PROPOSITIONS

The common assumption that the words "what S said" refer to a statement (or proposition) and that a statement so understood cannot reasonably be identified with a string of words understood in a certain way is dubious for reasons I have not mentioned in the text. The paper by Cartwright (1962) contains what is perhaps the classic defense of this assumption. In what follows I outline some of the considerations that make me doubt that the assumption is true.

Cartwright opened his discussion with a little dialogue between two people, A and B, who are discussing the French Defense, a standard strategy in chess.[15] Speaking of this defense, A says, "Botvinnik uses it," to which B replies, "That is true. But he lost with it against Tal." According to Cartwright, in uttering "That's true" B clearly predicates truth of something to which he *refers*. But exactly what is the object of B's reference? Although one might naturally suppose it is the words B hears, namely "Botvinnek uses it" (understood in the way A intended), Cartwright identifies it as "what A said (asserted, stated)" or "the statement A made" or "the statement that Botvinnek uses the French Defense." This identification is not philosophically sufficient, he adds, because although we know very well how to identify what B referred to, we may nevertheless "mistake other things for it." Cartwright discusses eight items that we might mistakenly identify with the statement in question. He is unable to say exactly what the statement is, but he argues that it cannot be one of these other things. His eight items includes asserting that p, uttering the words "Botvinnik uses it," the meaning of these words, tokens of the type, and four other things, but not the alternative I have just mentioned.

Cartwright's arguments relating to the statement B made consist largely of analytical remarks about the way we commonly speak. His analytical remarks do not, as he emphasized, provide a revealing picture of what he takes a statement (or proposition) to be, but others have relied on Cartwright's remarks in advancing a more explicit view of such a thing. Scott Soams is a case in point, and it will be useful to describe his view before continuing with Cartwright. According to Soames,[16] who refers to Cartwright's discussion with approval, "a statement…[is] that which someone states, not his stating of it" (p. 279). "To state that so and so is, of course," he adds, "not the same as to assume, believe, conjecture, or claim that so and so. However, what is stated can also be assumed, believed, conjectured, claimed, or asserted. Thus, statements can be identified with assumptions, beliefs, conjectures, claims, and assertions, provided that these terms are taken to refer to what is assumed, believed, conjectured, claimed, or asserted…. Since it is perfectly correct to speak of beliefs, assumptions, and the rest as being true, what might initially have seemed to be a multiplicity of different truth-bearers can be reduced substantially. In all of these cases, we predicate truth of propositions" (p. 280).

Later in his book Soames lists the central assumptions of an expository framework in which propositions as he understands them are featured. These assumptions include the following:

A1. Some things are asserted, believed, and known. Propositional attitudes like assertion, belief, and knowledge are relations that hold between agents and the things they assert, believe, and know.

---

[15] Cartwright (1962), pp. 83-84.
[16] Soams (2003).

A2. The things asserted, believed, and known may be expressed by sentences and designated by clauses such as ***the statement that S, the belief that S,*** … or simply ***that S***. I [Soames] will call the things designated by these clauses *propositions.*

A5. Propositional attitude ascriptions—***x asserts/believes/knows…that S***—report that an agent asserts, believes… [etc.] the proposition designated by ***that S*** (p. 373).

It is clear from Soames' words that he construes psychological verbs such as "believes" and "knows" as having proposition-denoting clauses as direct objects. Thus, he speaks freely of asserting, believing, and knowing propositions. But this usage is grammatically deviant. The point stands out when one considers a wider variety of psychological expressions. Consider the following: *asserting*, *doubting, fearing, suspecting, trusting, having no doubt, wondering,* and *wanting.* To my ear, locutions such as *knowing a person, believing a person, trusting a person, and fearing an enemy* are grammatically impeccable, but *knowing a proposition, believing a proposition, thinking a proposition, and asserting a proposition* sound odd, and *fearing a proposition, suspecting a proposition, trusting a proposition, having no doubt a proposition, hoping a proposition, wishing a proposition, deciding a proposition,* and *objecting a proposition* sound nonsensical. (Does a person who "fears a proposition" think the proposition is dangerous?) Soames might reply that these locutions make good sense if they are understood in the right way—if, say, they are understood as different ways of relating a person to proposition—but this reply is certainly ineffective if the technical notion of a proposition is to be accepted on the basis of "our ordinary ways of understanding ourselves and our words," as he implies in his criticism of Quine's views on language and meaning (see p. 281).

This brings me back to Cartwright, who is also explicitly concerned with everyday thinking about speech and reference. According to him, the person saying "That's true" referred to *what person A said* and also predicated truth of this entity, this statement, as Cartwright called it. But his assumption that phrases such as "what A said" are normally used to refer to objects that can also, at least as Soames claimed, be denoted by that-clauses is actually very dubious. To see this, suppose a person S said that P. Does "I know what S said" imply "I know that P"? Obviously not: "I know that P" implies that "P" is true, but people often speak falsely. Thus, when S said that P, the phrase "what S said" cannot be replaced by "that P" in the context "I know…." What can be put in place of "what S said" in this context is the clause "that S said that P," for if S said that P and I know what S said, I know that S said that P.

What explains this last implication? What is it about the meaning of "what S said" in the context "I know…" that justifies the conclusion that I have identified? The answer was given, I believe, by J.L. Austin in his famous paper, "Other Minds": The "what" following "I know" is not a relative pronoun equivalent to "that which" but an interrogative pronoun, like the Latin "*quid.*"[17] The sentence "I know what S said" has the sense, approximately, of "I know the answer to the question 'What did S say?'" If S said that P and I know this answer, I know that S said that P. Since "I know what S thinks (believes, suspects, etc.)" has a similar implication, it is a mistake to suppose that "what S thinks (believes, suspects, says)" *generally* denotes any object at all."

---

[17] Austin (1961).

What holds for "what S said, or says" in the context "I know…" does not, of course, hold for this clause in the context "I said…" or "I believe…," but the consequences of taking "what S said, or says" to denote an object in this context and of assuming that such an object is also denoted by a that-clause in sentences such as "I fear…," "I hope…," and "I have no doubt" make such a position highly questionable. Cartwright did not explicitly commit himself to this view of that-clauses, so he might possibly elude this objection. But most friends of propositions do accept this view, and they must come to terms with it.

How should we understand a that-clause that is attached to a psychological verb or verbal phrase? The easiest cases to understand contain verbs pertaining to overt speech. When an ordinary direct quotation follows "S said," as it does in my version of Cartwright's little dialogue, the quotation *exhibits* the words the speaker is supposed to have used in his/her speech act; in doing so it characterizes the semantic character of the speaker's act by means of those words and does not plausibly refer to any additional object: its semantic value is truth or falsity rather than something appropriate to a referring noun phrase. When a that-clause follows "S said…," the clause is what traditional grammarians called an "indirect quotation" of the speaker's words. As such, its semantic value corresponds to that of the sentence it indirectly quotes.[18] Parity of reasoning suggests that in the context "X said…," the clause "what S said" is a placeholder for a direct or indirect quotation rather than a designator for a propositional object that is not plausibly denoted by either kind of quotation. You can *hear* (literally hear) what a person says, and you can *see* (literally see) what a person writes on a blackboard.

In the final section of chapter four I discuss what I am calling direct and indirect quotation in more detail. To conclude this brief comment on the relation of "what X said" to propositions, I will simply observe that a direct quotation is actually the most obvious object of reference in the little dialogue that Cartwright presented. The person A is described as having said "Botvinnik uses it," and when B replies "That's true," the obvious object of his reference is not an elusive entity denoted by "what he said" but that quotation itself. If B had wanted to be fully explicit, he could have used the word "statement" as Quine does and then said: "Your statement, 'Botvinnik uses it,' is true." This candidate is *not* one of the eight items that Cartwright argues we could not rightly identify with the object of B's reference.

When a that-clause is introduced by a psychological verb such as "fears," "believes," "doubts," or "decides," the ingredient words have a more indirect function. To understand those words one must reflect on the meaning of the psychological verb. Consider the sentence "Harry fears that the bull will charge." To have this particular fear, Harry must be frightened, I should think, by the thought that the bull will charge. This thought would ideally be expressed in English by "The bull will charge" and ideally expressed in other languages by sentences that are best translated by "The bull will charge."[19] So in this case the words in the that-clause characterize a relevant thought by exhibiting words that would naturally be used in expressing it verbally. The that-clause in a sentence such as "Betty believes that snow is white" has a similar function: it exhibits words that would naturally be used in expressing the thought that snow is white. This thought is characteristic of the belief that snow is white because its occurrence is a central causal factor in the thinking and acting that a person having that belief is apt to engage in.[20]

---

[18] See chapter four, pp. 88-90.

[19] Qualifications are appropriate here. To understand what they are, see the penultimate section of chapter four, "Meaning, Intending, and Content Clauses."

[20] For further discussion, see the previous footnote.

Appendix 5

CHISHOLM'S DEFENSE OF CP

In view of his influence in epistemology, Chisholm's means of defending the principle CP deserves a brief comment. CP is the principle:

> If S accepts h and if h is not disconfirmed by S's total evidence, then h is probable for S.

Chisholm said that this principle can be justified by an a priori truth and a general presupposition that epistemologists make when they ask Socratic questions about evidence and justification. The presupposition is best stated first:

> I am justified in believing that I can improve and correct my system of beliefs. Of those that are about matters of interest and concern to me, I can eliminate the ones that are unjustified and add others that are justified; and I can replace less justified beliefs about those topics by beliefs that are more justified.

The a priori truth needed for Chisholm's justification of CP is a conditional statement whose antecedent is the conjunction of the propositions just given and whose consequent is the principle to be justified. If G expresses the general presupposition as a single complex statement, then G → CP is an entailment that can, according to Chisholm, be known to be true a priori.

If the statement G → CP is in fact true, then if G is true, the principle CP is also true. Chisholm made no effort to prove that either of these premises is true, however. He simply asserted that the conditional statement is true and said that G expresses "that faith in oneself with which the epistemologist sets out." Because the required connection between G and CP is exceedingly hard to see, I seriously doubt that the conditional G → CP is actually true. Yet even if it is true, Chisholm's justification for CP remains highly questionable for reasons that deserves to be pointed out.

Since Chisholm advances the assertion G as an expression of faith, the operative premise in his argument should really be represented by F(G), where F is a sentential operator meaning "I have faith that," the pronoun referring to any epistemologist who may rehearse the argument. The two premises F(G) and G → CP obviously do not entail CP or provide any evidence that CP is true. Do they entail F(CP)? I would say no; the operator F is not closed under entailment. Is it closed under known logical entailment? That is, if someone knows that G entails CP, will he or she inevitably have faith in CP? Possibly so. But the conclusion would only be another statement of faith, F(CP), not CP itself. No doubt Chisholm believes what he says. But a justification for CP should amount to more than this.

Appendix 6

ANALYTIC PROBABILITY PRINCIPLES

The version of the probability calculus that I interpret in relation to the certainty and evidential support of statements contains just three axioms and one basic definition of conditional probability. The first axiom identifies the range of values

appropriate to a categorical probability statement. Each such statement is stipulated to have a unique certainty value (or epistemic probability) in the domain of real numbers between 0 and 1.

A1. $0 \leq P(p) \leq 1$.

According to the second axiom, a statement that is certainly true (or certain) has the probability value of 1. Formally,

A2. If p is certain (L-true or analytic), $P(p) = 1$.

The third and final axiom specifies a certainty value for disjunctions consisting of propositions whose mutual incompatibility is certain. Specifically:

A3. If it is certain that p and q are mutually exclusive—if it is certain, that is, that $p \supset \sim q$ and, therefore, certain that $q \supset \sim p$—then $P(p \vee q) = P(p) + P(q)$.[21]

The following definition completes the basis for this system of epistemic probability. It introduces the concept of conditional epistemic probability, which is understood as a measure of evidential support between statements.

D1. $$P(p/q) = \frac{P(p \wedge q)}{P(q)}, \text{ if } P(q) \neq 0.$$

The definition may not seem intuitive at first, but it works. Like any other definition, it is in principle arbitrary and yields an analytic truth according to the ideas set forth in chapter four.

[21] Since statements that are mutually exclusive are not necessarily contradictory, the value of a disjunction involving them may be less than 1. Thus, Tom's being six foot tall is incompatible with his being five foot tall, but he may be neither. In addition to being mutually exclusive, contradictory statements are jointly exhaustive: one or the other is always true.

References

Adams, Ernest. 1988. "Modus Tollens Revisited," *Analysis,* 28: 122-127.

Allaire, Edwin. 1963. "Bare Particulars," *Philosophical* Studies, XVI:

Almog, Joseph. 2005. "Is a Unified Description of Language-and-Thought Possible?", *The Journal of Philosophy*, 111, 10:493-531.

Amis, Kingsley. 1997. *The King's English*. New York: St. Martin's Press.

Armstrong, D. M. 1978. *Universals and Scientific Realism*, 2 vols. Cambridge, UK: Cambridge University Press.

Armstrong, D. M. 1984. "[Reply] To Aune," in *D.M. Armstrong*, ed. Radu J. Bogdan. Dordrecht: Reidel. Pp. 250-256.

Armstrong, D. M. 2004 (A). *Truth and Truthmaking*. Cambridge, UK: Cambridge University Press. Pp. 139-154.

Armstrong, D.M. 2004 (B). "How do Particulars Stand to Universals," in in Dean W. Zimmerman, ed., *Oxford Studies in Metaphysics,* vol. 1. Oxford: Clarendon Press. Pp.

Anscombe, G.E.M. 1981. *An Introduction to Wittgenstein's Tractatus*. London: Hutchinson.

Augustine. 1951. *The Confessions of Saint Augustine*, trans. Edward B. Pusey. New York: Pocket Books.

Audi. Robert. 1999. "Self-Evidence," in James E. Tomberlin, ed., *Philosophical Perspectives*, Vol. 13. Cambridge, MA: Blackwell. Pp. 205-228.

Aune, Bruce. 1967. *Knowledge, Mind, and Nature.* Atascadero, CA: Ridgeview Publishing Company.

Aune, Bruce. 1970. *Rationalism, Empiricism, and Pragmatism*. Atascadero, CA: Ridgeview.

Aune, Bruce. 1978. *Reason and Action*. Dordrecht, Holland: Reidel.

Aune, Bruce. 1984. "Armstrong on Universals and Particulars," in R.J. Bogden, *D.M. Armstrong*. Dordrecht, Holland: Reidel. Pp. 161-170.

Aune, Bruce. 1985. *Metaphysics: the Elements.* Minneapolis: University of Minnesota Press.

Aune, Bruce. 1990. "Action, Inference, Belief, and Intention," in James Tomberlin, ed., *Philosophical Perspectives*, vol. 4. Atascadero, CA: Ridgeview. Pp. 247-272.

Aune, Bruce. 1991. *Knowledge of the External World*. London and New York: Routledge.

Aune, Bruce. 2002. "Against Moderate Rationalism," *Journal of Philosophical Research*, 27 (2002), 1-26.

Aune, Bruce. 2002. "Universals and Predication," in Richard Gale, ed., *The Blackwell Guide to Metaphysics*. Oxford: Blackwell. Pp. 131-150.

Austin, J. L. 1961. *Philosophical Papers.* Oxford: Clarendon Press. Pp. 123-152.

Austin, J. L. 1962. *How To Do Things With Words.* Oxford: Clarendon Press.

Ayer, A. J. 1956. *The Problem of Knowledge*. Baltimore, MD: Penguin Books.

Barnes, W.H.F. 1944-5. "The Myth of Sense Data," *Proceedings of the Aristotelian Society*, XLV; reprinted in Robert J. Swartz (1965), pp. 138-167.

Bach, Kent. 2005. "The Emperor's New 'Knows'," in Gerhard Preyer and Georg Peter, eds., *Contextualism in Philosophy*. Oxford: Clarendon Press. Pp. 51-90.

Baxter, Donald L. M. 2001. "Instantiation as Partial Identity," *Australasian Jourtnal of Philosophy*, 79: 449-464.

Bealer, George. 1992. "The Incoherence of Empiricism," *The Aristotelian Society, Supplementary Volume,* 66: 99-138.

Bealer, George. 1999a. "A Theory of the A Priori," *Philosophical Perspectives*, vol. 13. Atascadero, CA: Ridgeview. Pp. 30-55.

Bealer, George. 1999b. "The A Priori," in John Greco and Ernest Sosa, eds, *The Blackwell Guide to Epistemology*. Malden, MA: Blackwell. Pp. 243-270.
Belloc, Hilaire. 1970. *Complete Verse.* London: Duckworth.
Bogden, Radu J. 1984. *D. M. Armstrong.* Dordrecht: Reidel. Pp.
Boghossian, Paul. 2000. "Knowledge of Logic," in Paul Boghossian and Christopher Peacock, eds. *New Essays on the A Prior.* Oxford: Clarendon Press. Pp. 229-254.
BonJour, Laurence. 1998. *In Defense of Pure Reason*. Cambridge: Cambridge University Press.
BonJour, Laurence. 1999. "Foundationalism and the External World," in *Philosophical Perspectives*, 13. Pp. 229-249.
BonJour, Laurence. 2004. "In Search of Direct Realism," *Philosophy and Phenomenological Research,* 69, 2: 349-367.
BonJour, Laurence. 2005. "In Defense of the A Priori," in Mattias Steup and Ernest Sosa, eds., *Contemporary Debates in Epistemology*. Malden, MA: Blackwell. Pp. 98-104.
BonJour, Laurence. 2007. "Epistemological Problems of Perception"*, The Stanford Encyclopedia of Philosophy* (Summer 2007 Edition), Edward N. Zalta, ed. *URL = <http://plato.stanford.edu/archives/sum2007/entries/perception-episprob/>.*
Breuckner, Tony. 2004. "Brains in a Vat", *The Stanford Encyclopedia of Philosophy* (Winter 2004 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2004/entries/brain-vat/>.
Burge, Tyler. 1979. "Individualism and the Mental," in *Midwest Studies in Philosophy*, vol. 4. Pp. 73-121.
Butler, Joseph. 1939. *The Whole Works of Joseph Butler, LL.D.* London: Thomas Tegg.
Byrne, Alex and David R. Hilbert. 1997. *Readings on Color*, vol. 2: *The Science of Color*. Cambridge, MA: MIT Press.
Campbell, Keith. 1997. "The Metaphysics of Abstract Particulars," in D.H. Mellor and Alex Oliver, eds., *Properties*. Oxford: Oxford University Press. Pp. 125-139.
Carnap, Rudolf. 1928. *Der Logische Aufbau der Welt. Berlin: Weltkreis.* Translated by R. George as *The Logical Structure of the World*. Berkeley and Los Angeles: University of California Press, 1967.
Carnap, Rudolf. 1936. "Testability and Meaning," *Philosophy of Science,* 3 and 4 (1937); reprinted, with omissions, in *Readings in The Philosophy of Science*, Herbert Feigl and May Brodbeck, eds. New York: Appleton Century Crofts, 1953. Pp. 47-92
Carnap, Rudolf. 1952b. "Meaning Postulates," *Philosophical Studies*, 3: 65-73. Reprinted in Carnap (1956), pp. 222-229.
Carnap, Rudolf. 1956a. *Meaning and Necessity*, 2nd ed. Chicago: University of Chicago Press.
Carnap, Rudolf. 1956b. "The Methodological Character of Theoretical Concepts," in Herbert Feigl and Michael Scriven, eds., *Minnesota Studies in the Philosophy of Science*, Vol. 1. Minneapolis: University of Minnesota Press. Pp. 38-76.
Carnap, Rudolf. 1966. *Philosophical Foundations of Physics*. New York: Basic Books.
Carnap, Rudolf. 1958. *Introduction to Symbolic Logic and Its Applications*, trans. W.H. Meyer and John Wilkinson. New York: Dover.
Carnap, Rudolf. 1990. "Quine on Analyticity," in Richard Creath, ed., *Dear Carnap, Dear Van.* Berkeley: Univ. of Cal. Press. Pp. 427-432.

Cartwright, Richard. "Propositions," in Ronald J. Butler, ed., *Analytical Philosophy.* Oxford: Blackwell. Pp. 81-103.
Chisholm, R. M. 1976. *Person and Object*. London: Allen and Unwin.
Chisholm, R.M. 1979. "Objects and Persons: Objections and Replies," *Grazer Philosophische Studien*, 7/8: 317-388.
Chisholm, R.M. 1996. *Theory of Knowledge*, 3rd ed. Englewood Cliffs, NJ: Prentice Hall.
Chisholm, R. M. 1997. *The Philosophy of Roderick Chisholm*, L. E. Hahn, ed. Chicago & La Salle: Open Court.
Chisholm, R. M. 1991. "On the Simplicity of the Soul," in James E. Tomberlin, ed., *Philosophical Perspectives*, Vol. 5. Atascadero: Ridgeview Publishing Company. Pp. 167-181.
Church, Alonzo. 1936. "A Note on the Enscheidungsproblem," *Journal of Symbolic Logic*, 1:40f, 101f.
Church, Alonzo. 1956. *An Introduction to Mathematical Logic.* Princeton, NJ: Princeton University Press.
Cohen, Stuart. 2000. "Contextualism and Skepticism," in *Skepticism*, ed. Ernest Sosa and Enrique Villaneuva. Boston, MA & Oxford, UK: Blackwell.
Crane, Tim, "The Problem of Perception", *The Stanford Encyclopedia of Philosophy* (Winter 2006 Edition), Edward N. Zalta ,ed. URL = http://plato.stanford.edu/archives/win2006/entries/perception-problem/>.
Creath, Richard (ed.). 1952. *Dear Carnap, Dear Van.* [City: pub]
Davidson, Donald. 1986. "A Coherence Theory of Truth and Knowledge," in Ernest LePore, ed., *Truth and Interpretation*. Oxford: Blackwell. Pp. 307-319.
DeRose, Keith. 1995. "Solving the Skeptical Problem," *Philosophical Review*, 104:1-52.
DeRose, Keith and Ted Warfield (eds). 1999. *Skepticism: A Contemporary Reader*. New York: Oxford University Press.
Descartes, René. 1985. *Principles of Philosophy*, in *The Philosophical Writings of Descartes,* Vol. 1, John Cottingham, Robert Soothoff, and Dogald Murdoch, eds. and trans. Cambridge, UK: Cambridge University Press, 1984.
Dretske, Fred. 1970. "Epistemic Operators," Journal of Philosophy, 67: 1007-1023.
Dretske, Fred. 1982. *Knowledge and the Flow of Information.* Cambridge, MA: MIT Press.
Dretske, Fred. 2005. "The Case Against Closure," in Mattias Steup and Ernest Sosa, eds., *Contemporary Debates in Epistemology*. Malden, MA: Blackwell. Pp. 13-25.
Earman, John (ed). 1983. *Testing Scientific Theories*, vol. X of *Minnesota Studies in the Philosophy of Science.* Minneapolis: University of Minnesota Press.
Feigl, Herbert. 1943. "Logical Empiricism," in D.D. Runes, ed. *Twentieth Century Philosophy*. New York: Philosophical Library. Reprinted with omissions in Feigl and Sellars (1949), pp. 3-26.
Feigl, Herbert and Wilfrid Sellars (eds.). 1949. *Readings in Philosophical Analysis.* New York: Appleton-Century-Crofts.
Feldman, Richard. 1999. "Contextualism and Skepticism," in James Tomberlin, ed., *Philosophical Perspectives, vol. 13. Cambridge, MA: Blackwell. Pp. 91-114.*
Feldman, Richard. 2005. "Justification is Internal," in Mattias Steup and Ernest Sosa, eds., *Contemporary Debates in Epistemology*. Malden, MA: Blackwell. Pp. 270-284.
Ferguson, Niall. 2002. *Empire*. New York: Basic Books.
Fine, Kit. 2005. "Our Knowledge of Mathematical Objects," in Tamar Szabó Gendler and John Hawthorne, eds., *Oxford Studies in Epistemology*, vol. 1. Oxford: Clarendon Press. Pp. 89-110.

Fitch, Greg. 2002. "Singular Propositions," *The Stanford Encyclopedia of Philosophy* (Summer 2002 Edition), Edward N. Zalta (ed.), URL = http://plato.stanford.edu/archives/summer 2002/entries/propositions-singular/>.
Fitch, Gregory. *Saul Kripke.* 2004. Montreal: McGill-Queeen's University Press.
Fodor, Jerry. 2000. *The Mind Doesn't Work That Way.* Cambridge, MA: MIT Press.
Fogelin, Robert J. 1994. *Pyrrhonian Reflections on Knowledge and Justification*. New York: Oxford University Press.
Fogelin, Robert. 2000. "Contextualism and Externalism: Trading One Form of Skepticism for Another," in Ernest Sosa and Enrique Villaneuva, eds., *Philosophical Issues, 10: Skepticism. Boston, MA: Blackwell. Pp. 43-57.*
Frege, Gottlob. 1892. "On Concept and Object," trans. P.T. Geach in Frege (1950), pp. 42-55.
Frege, Gottlob. 1950. *Foundations of Arithmetic* second ed., trans. J. L. Austin. Oxford: Blackwell. (Originally published in German in 1884.)
Frege, Gottlob. *Translations from the Philosophical Writings of Gottlob Frege*, ed. Peter Geach and Max Black. Oxford: Blackwell, 1952.
Frege, Gottlob. 1962. "Compound Thoughts," trans. R.H. Stoothoff, *Mind*, LXXII, 1-17; reprinted in Klempke (1968), 537-558.
Frege, Gottlob. 1965. "The Thought: A Logical Inquiry," trans. A.M. and Marcelle Quinton. Mind, LXV: 289-311. (Originally published 1918-19 in *Beiträge zur Philosophie des Deutschen Idealismus*,)
Freidman, Michael. *Reconsidering Logical Positivism*. Cambridge, UK: Cambridge University Press. 1999.
Fumerton, Richard. 2005. "The Challenge of Refuting Skepticism," in Mattias Steup and Ernest Sosa, eds., *Contempoary Debates in Epistemology*. Malden, MA: Blackwell. Pp. 85-97.
Gauker, Christopher. 1994*. Thinking Out Loud. Princeton:* Princeton University Press.
Geach, P. T. 1969. *God and the Soul*. London: Routledge & Kegan Paul.
Gettier, Edmund. 1963. "Is Justified True Belief Knowledge?" *Analysis,* 23:121-123.
Gödel, Kurt. 1951. "Russell's Mathematical Logic," in Paul A. Schilpp, ed., *The Philosophy of Bertrand Russell*. New York: Tudor Publishing Company. Pp. 125-153.
Goldman, Alvin. 1976. "Discrimination and Perceptual Knowledge," *Journal of Philosophy, 73:771-791.*
Goldman, Alvin. 1979. "Reliability: What is Justified Belief?" in G.S. Pappas, ed., *Justification and Knowledge*. Reidel: Dordrecht. Pp. 1-23.
Goodman, Nelson. 1951. *The Structure of Appearance*. Indianapolis: Bobbs-Merrill.
Geach, P. T. 1956. *Mental Acts*. London: Hutchinson.
Gribben, John, and Jeremy Cherfas. 1983. *The Monkey Puzzle*. New York: McGraw-Hill.
Gribben, John. 1998. *Almost Everyone's Guide to Science.* Yale: New Haven, CT.
Grunstra, Bernard R. 1969. "The Plausibility of the Entrenchment Concept," in Nicholas Rescher, ed., *Studies in the Philosophy of Science.* Oxford, UK: Blackwell. Pp. 100-127.
Hacking, Ian. 2005. "Why Race Still Matters," *Daedalus. Winter issue. Pp. 102-116.*
Hájek, Alan and Brian Skyrms. 2000. Review of *Bayes or Bust?* By John Earman. *Philosophy and Phenomenological Research*, LX, No. 3: 707-711.
Harmon, Gilbert. 1996. "Analyticity Regained?" *Noûs,* 30, 3: 392-400.
Hawthorne, John. 2004. *Knowledge and Lotteries.* Oxford: Clarendon Press.

Hawthorne, John. 2005. "The Case for Closure," in Mattias Steup and Ernest Sosa, eds., *Contemporary Debates in Epistemology*. Malden, MA: Blackwell. Pp. 26-43.

Heil, John. 2004. "Properties and Powers," in Dear W. Zimmerman, ed., *Oxford Studies in Metaphysics,* 223-254.

Heller, Mark. 1999. "The Proper Role for Contextualism in an Anti-luck Epistemology," in James E. Tomberlin, ed., *Philosophical Perspectives*, vol, 13. Cambridge, MA: Blackwell. Pp. 115-129.

Huemer, Michael. 2007. "Sense-Data", *The Stanford Encyclopedia of Philosophy* (Fall 2007 Edition), Edward N. Zalta, ed. URL = <http://plato.stanford.edu/archives/fall2007/entries/sense-data/>.

Hintikka, Jaakko. 1999. "The Emperor's New Intuitions," *Journal of Philosophy*, Vol. 92, No. 3 (1999): 127-147.

Howson, Colin and Peter Urbach. 1989. *Scientific Reasoning: The Bayesian Approach.* Indianapolis: Bobbs-Merrill.

Horwich, Paul. 2000. "Stipulation, Meaning, and Aprioricity," in Paul Boghossian and Christopher Peacock, eds. *New Essays on the A Prior.* Oxford: Clarendon Press. Pp. 150-169.

Hume, David. 1978. *A Treatise of Human Nature*, second ed., L. A. Selby-Bigge and P. H. Niddich, eds. Oxford: Clarendon Press. (The *Treatise* was originally published in 1739.)

Hume, David. 1902. E*nquiry Concerning the Human Understanding,* in L. A. Selby-Bigge, ed, *Hume's Inquiries*. Oxford: Clarendon Press. (The *Enquiry* was originally published in 1777.)

Jackson, Frank. 2000. "Representation, Scepticism, and the A Priori," in Paul Boghossian and Christopher Peacock, eds., *New Essays in the A Priori.* Oxford: Clarendon Press, 2000. Pp. 320-332.

Jackson, Frank and Michael Smith (eds.). 2005. *The Oxford Handbook of Contemporary Philosophy*. Oxford, UK: Oxford University Press.

Johnson, Mark. 2006. "Better than Mere Knowledge? The Function of sensory Awareness," in Tamar Szabó Gendler and John hawthorne, eds., *Perceptual Experience.* Oxford: Clarendon Press.

Kant, Emmanuel. 1974. *Logic*, trans. Robert Hartman and Wolfgang Schwartz. Indianapolis, IA: Bobbs-Merrill. (German original first published in 1800.)

Kant, Emmanuel. 1997. *Critique of Pure Reason*, trans. and ed. Paul Guyer and Allen Wood. Cambridge, UK: Cambridge University Press. (German original first published in 1781.)

Kaplan, David. 1971. "Homage to Rudolf Carnap," in ," *PSA 1970*, vol. 8 of *Boston Studies in the Philosophy of Science*, ed. Roger C. Buck and Robert S. Cohen. Dordrecht: Reidel. Pp. XLV-XLVII.

Kaplan, David. 1991a. "Dthat, in *Meaning and Truth: the Essential readings in Modern Semantics*, ed. Jay Garfield and Murray Kiteley. New York: Paragon House. Pp. 212-233. (This essay was originally written and circulated in 1970.)

Kaplan, David. 1992a. "Demonstratives," in *Themes from Kaplan*, ed. J. Almog, J. Perry, and H. Wetstein. New York: Oxford University Press. Pp. 481-563. (This essay was originally written in two parts and circulated in the early 1970s.)

Kaplan, David. 1992b. "Afterthoughts," in *Themes from Kaplan*, ed. J. Almog, J. Perry, and H. Wetstein. New York: Oxford University Press. Pp. 565-614. (This essay was originally written in 1988.)

Katz, Bernard. 1999. "On a Supposed Counterexample to *Modus ponens*," *Journal of Philosophy*, XCVI: 404-415.

Kay, Paul and Chad K. McDaniel. 1997. "The Linguistic Significance of the Meanings of Basic Color Terms," in Alex Byrne and David R. Hilbert, eds., *Readings on Color*, Vol. 2: *The Science of Color.* Cambridge, MA: MIT Press, pp. 399-442.

Kemp Smith, Norman. 1948. *Hume"s DIALOGUES CONCERNING NATURAL RELIGION*. New York: Social Scienecs Publishers. (Originally published in 1779.)

King, Jeffrey. 2001. "Structured Propositions," *Stanford Encyclopedia of Philosophy* (Fall 2002 Edition), Edward N. Zalta (ed.), URL = http://plato.stanford.edu/archives/fall 2001/entries/propositions-structured/>.

Klempke, E.D. 1968. *Essays on* Frege. Urbana, IL: University of Illinois Press.

Kneale, William and Martha. 1962. *The Development of Logic.* Oxford: Clarendon Press.

Kornblith, Hilary. 1999. "Knowledge in Humans and Other Animals," in James Tomberlion, ed., *Philosophical Perspectives*, vol. 13. Oxford: Blackwell Publishers. P. 327-346.

Körner, Stephan. 1960. *The Philosophy of Mathematics.* London: Hutchinson.

Kripke, Saul. 1979. "A Puzzle about Belief," in Avishai Margalit, ed., *Meaning and Use*. Reidel: Dordrecht. Pp.

Kripke, Saul. 1980. *Naming and Necessity*, rev. ed. Cambridge, MA: Harvard University Press.

LaPorte, Joseph. *Natural Kinds and Conceptual Change.* Cambridge: Cambridge University Press.

Lehrer, Keith. 1997. *Self Trust*. Oxford: Clarendon Press.

Lewis, David. 1973. *Counterfactuals*. Oxford: Blackwell, 19

Lewis, David. 1983a. "New Work for a Theory of Universals," *Australasian Journal of Philosophy*, 61/4: 343-77. Reprinted in Lewis (1999). Pp. 8-55.

Lewis, David. 1983b. "Scorekeeping in a Language Game," in Lewis, *Philosophical Papers,* Vol. 1. Oxford: Oxford University Press. Pp. 233-249.

Lewis, David. 1986. *On the Plurality of Worlds.* Oxford: Blackwell.

Lewis, David. 1996. "Elusive Knowledge," *Australasian Journal of Philosophy*, 74 549-567. Reprinted in Lewis (1999). Pp. 418-446.

Lewis, David. 1999. *Papers in Metaphysics and Epistemology*. Cambridge, UK: Cambridge University Press.

Liddell, H.G. and Scott. 1968. *Intermediate Greek-English Lexicon*. Oxford: Clarendon Press.

Lipton, Peter. 2004. *Inference to the Best Explanation*. London: Routledge.

Locke, John. 1984. *An Essay Concerning Human Understanding*, 2 vols., ed. G. A. Fraser. Oxford: Clarendon Press.

Ludlow, Peter. "Contextualism and the New Linguistic Turn in Epistemology," in Gerhard Preyer and Georg Peter, eds., *Contextualism in Philosophy*. Oxford: Clarendon Press. Pp.11-50.

Lycan, William G. 1988. *Judgment and Justification*. Cambridge: Cambridge University Press.

Macfarlane, John. 2005. "The Assessment Sensitivity of Knowledge Ascriptions," in Tamar S. Gendler and John Hawthorne, eds., *Oxford Studies in Epistemology*, vol. 1. Oxford: Clarendon Press. Pp. 197-233.

Margolis, Eric and Stephen Laurence. 1999. *Concepts: Core Readings*. Cambridge, MA: MIT Press.

Martin, R. M. 1958. *Truth and Denotation.* Chicago: University of Chicago Press.

Mates, Benson. 1986. *The Philosophy of Leibniz*. Oxford: Oxford University Press.

McGee, Vann. 1985. "A Counterexample to *Modus Ponens*," *Journal of Philosophy*, 82: 462-471.

McGinn, Colin. 2002. *The Making of a Philosopher.* New York: HarperCollins.

Mill, John Stuart. 1861. *Utilitarianism.*
Mill, John Stuart. 1859. *On Liberty.*
Moore, G. E. 1903. *Principia Ethics.* Cambridge, UK: Cambridge University Press.
Moore, G.E. 1969. "Certainty," in Moore, *Philosophical Papers*. London: All and Unwin. Pp. 227-251.
Nichols, Shaun, Stephen Stich, and Jonathan Weinberg. 2003. "Meta Skepticism: Meditations on Ethno-Epistemology," in S. Luper, ed., *The Skeptics.* Aldershot, UK: Newgate Publishing. Pp. 227-247.
Nisbett, Richard E. and Timothy D. Wilson. 1977. "Telling More than We Know: Verbal Reports on Mental Processes," *Psychological Review*, 84: 231-259.
Phillips, Lawrence. 1973. *Bayesian Statistics for Social Scientists*. London: Nelson.
Peacock, Christopher. "Explaining the A Priori: The Programme of Moderate Rationalism," in Paul Boghossian and Christopher Peacock, eds. *New Essays on the A Priori.* Oxford: Clarendon Press. Pp. 255-285.
Perry, John. 1979. "The Problem of the Essential Indexical," *Nous,* 13/1: 3-21.
Pinker, Steven 1994. *The Language Instinct.* New York: Harper-Collins.
Pinker, Steven. 1999. *Words and Rules*. New York: Basic Books.
Plantinga, Alvin. 1990. "Justification in the 20th Century," reprinted in Pojman (1999). Pp. 357-373.
Plantinga, Alvin. 1993. *Warrant and Proper Function*. New York: Oxford University Press.
Pojman,Louis P. 1999. *The Theory of Knowledge,* 2nd ed. Belmont, CA: Wadsworth.
Priest, Graham. 1998. "What Is So Bad About Contradictions?" *Journal of Philosophy*, XCV, Number 8: 410-426
Prior, A. N. 1960. "The Runaway Inference Ticket," *Analysis* 21 (1960), 38-9.
Putnam, Hilary. 1981. *Reason, Truth, and History.* Cambridge, UK: Cambridge University Press.
Quine, W.V.O. 1953. "Two Dogmas of Empiricism," in Quine, *From a Logical Point of View* . Cambridge, MA: Harvard University Press. Pp. 20-46.
Quine, W.V.O. 1956. *Mathematical Logic.* Cambridge, MA: Harvard University Press.
Quine, W.V.O. 1960. *Word and Object*. Cambridge, MA: MIT Press.
Quine, W.V.O. 1969. *Set Theory and its Logic*, rev. ed. Cambridge, MA: Harvard University Press.
Quine, W.V.O. 1981. "Success and Limits of Mathematization," in Quine, *Theories and Things*. Cambridge, MA: Harvard University Press. Pp. 148-155.
Quine, W.V.O. 1991. "Two Dogmas in Retrospect," *Canadian Journal of Philosophy*, XXI No. 3: 265-274
Quinton, Anthony. "The Problem of Perception," *Mind* (1965). Reprinted in Swartz (1965), pp. 497-526.
Ramsey, F. P. 1931. "Truth and Probability', in *The Foundations of Mathematics and Other Logical Essays*. London: Routledge & Kegan Paul. Pp. 156-198.
Reichenbach, Maria. 1970. "Homage to Carnap," in *PSA 1970*, ed. Roger C. Buck and Robert S. Cohen. Dordrecht: Reidel.
Rosenberg, Jay F. 2002. *Thinking About Knowing*. Oxford: Clarendon Press.
Rosenkrantz, Roger. 1981. *Foundations and Applications of Inductive Probability.* Atascadero, CA: Ridgeview Publishing Company.
Rosenkrantz, Roger. 1982. "Does the Philosophy of Induction Rest on a Mistake?", *Journal of Philosophy*, 79, 2 (1982): 78-97.
Russell, Bertrand. 1903. *The Principles of Mathematics.* Cambridge, UK: Cambridge University Press.

Russell, Bertrand. 1914. *Our Knowledge of the External World*. London: Allen & Unwin.
Russell, Bertrand. 1927. *Analysis of Matter.* London: Kegan Paul.
Russell, Bertrand. 1940. *An Inquiry Into Meaning and Truth.* London: Allen & Unwin.
Russell, Bertrand. 1948. *Human Knowledge: Its Scope and Limits*. London: Allen & Unwin.
Russell, Bertrand. 1951. "John Stuart Mill," in *Portraits from Memory.* New York: Simon and Schuster. Pp. 122-144.
Russell, Bertrand. 1953. "Knowledge by Acquaintance," in Russell, *Mysticism and Logic.* Harmonsworth, UK: Penguin. Pp. 197-218.
Salmon, Nathan U. 2005. *Reference and Essence*, second ed. Amherst, NY: Prometheus Books.
Salmon, Wesley. 1979. *The Foundations of Scientific Inference*. Pittsburgh, PA: University of Pittsburgh Press.
Schaffer, Jonathan. 2005. "Contrastive Knowledge," in Tamar Szabó Gendler and John Hawthorne, eds., *Oxford Studies in Philosophy*, vol. 1. Oxford: Clarendon Press. Pp. 235-271.
Schiffer, Stephen. "Interest-Relative Invariantism, *Philosophy and Phenomenological Research,* forthcoming.
Sellars, Wilfrid. 1959. "Empiricism and the Philosophy of Mind," in Herbert Feigl and Michael Scriven, eds., *Minnesota Studies in the Philosophy of Science*, vol. I. Minneapolis: University of Minnesota Press. Pp. 253-3329.
Sellars, Wilfrid. 1979. *Naturalism and Ontology*. Atascadero, CA: Ridgeview.
Shope, Robert K. 1983. *The Analysis of Knowing*. Princeton, NJ: Princeton University Press.
Sinnot-Armstrong, Walter, James Moore and Robert Fogelin. 1990. "A Defense of Modus Tollens," *Analysis*, 50: 9-16.
Sorensen, Roy, "Vagueness", *The Stanford Encyclopedia of Philosophy* (Fall 2006 Edition), Edward N. Zalta, ed. URL = <http://plato.stanford.edu/archives/fall2006/entries/vagueness/>.
Skyrms, Brian. 1986 *Choice and Chance: An Introduction to Inductive Logic*, third edition. Encino, CA: Dickinson Publishing Company.
Smyth, H. W. 1920. *Greek Grammar*. Cambridge, MA: Harvard University Press.
Stanley, Jason. 2005. *Knowledge and Practical Interests*. Oxford: Claendon Press.
Steup, Matthias. 1996. *An Introduction to Contemporary Epistemology*. Upper Saddle River, NJ: Prentice Hall.
Strawson, P.F. and H.P. Grice. 1956. "In Defense of Dogma," *Philosophical Review,* LXV, 141-158.
Suppes, Patrick. 1957. *Introduction to Logic.* Princeton, NJ: Van Nostrand.
Suppes, Patrick. 1960. *Axiomatic Set Theory*. Princeton, NJ: Van Nostrand.
Swartz, Robert J. 1965. *Perceiving, Sensing, and Knowing.* Garden City, NY: Anchor Books.
Talbot, William. 2001. "Bayesian Epistemology," *Stanford Encyclopedia of Philosophy* (Fall, 2001), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/Fall 2001/entries/bayesian epistemology/.>
Tarski, Alfred. 1949. "The Semantic Conception of Truth," in Herbert Feigl and Wilfrid Sellars, eds., *Readings in Philosophical Analysis.* New York: Appleton-Century-Crofts. Pp. 52-84.
Trehub, Arnold. 1991. *The Cognitive Brain.* Cambridge, MA: MIT Press.
Tversky, A. and D. Kahneman. 1983. ''Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment'', *Psychological Review* 91, 293–315.

Unger, Peter. 1975. *Ignorance: A Case for Skepticism.* Oxford: Oxford University Press.
Van Inwagen, Peter. 2004. "A Theory of Properties,' in Dean W. Zimmerman, ed., *Oxford Studies in Metaphysics,* vol. 1. Oxford: Clarendon Press. Pp. 107-138.
van Fraassen, Bas C. 1966. "Singular Terms, Truth-value Gaps, and Free Logic," *Journal of Philosophy*, 63: 481-95.
van Fraassen, Bas C. 1980. *The Scientific Image.* Oxford: Clarendon Press.
van Fraassen, Bas C. 1989. *Laws and Symmetry*. Oxford: Clarendon Press.
van Fraassen, Bas C. 2002. *The Empirical Stance.* New Haven: Yale University Press.
Van Fraassen, Bas C, James Ladyman, Igor Douven and Peter Lipton. 2006. "Wouldn't it be Lovely," *Metascience*, 14, 3 (2006): 331-361.
van Heijenoort, Jean. 1967. "Gödel's Theorem," *Encyclopedia of Philosophy*, Vol. 3. New York: Macmillan. Pp. 348-357.
Wetzel, Linda. 2000. "The Trouble with Nominalism," *Philosophical Studies*, XCVIII: 361-370.
Wetzel, Linda. 2003. *On Types and Tokens*. Cambridge, MA: MIT Press.
Williamson, Timothy. 1994. *Vagueness*. London: Routledge.
Williamson, Timothy. 2000. *Knowledge and Its Limits.* Oxford: Oxford University Press.
Wittgenstein, Ludwig. 1953. *Philosophical Investigations*, tr. G.E.M. Anscombe. Oxford: Blackwell.

Index

Adam, James, 93Allaire, Edwin, 73, 74, 153
Almog, Joseph, 89, 90, 153
analysis
    non-reductive, 60
    and explication, 10, 59, 60, 62, 68, 137, 145.
analytic/synthetic distinction, 47-70
analytic truth
    as understood by Carnap, 58
    Boghossian on, 146-149
    criticized by Quine, 47-53
    defended, 52-63
    broad sense of, 60
    relativized to person and time, 60
    doctrine extended, 57-62
Armstrong, D.M., iii, 16,35,71,72,73,74,75,76,77,82,146,147,153,154,160
    older theory of universals, 71-76
    new theory of universals 145
    on predicate nominalism, p. 74 note 21
Audi, Robert, vi, 153
Anscombe, Elizabeth, 81, 153
Aristotle, 3, 66, 71, 99, 100, 120, 124
    regress argument, 99, 100
Augustine, St., 118
Aune, Bruce, 71, 72, 74, 76, 83, 84, 87, 89, 95, 98, 105, 106, 113, 117, 127, 132, 138, 153
Austin, J.L., 87, 92, 155
Ayer, A.J., 98

Bach, Kent, 4, 5, 8, 9, 13, 153
Baker, Lynne, vii
Barnes, W.H.F., 98
Bayesians
    subjective, 132
Bayes' theorem, vii
    inferences based on, 128
    simple form of, 129
    general discussion of, 128-137.
Bealer, George, vi, 28, 39, 153
Belloc, Hillaire, 83
Bernays, Paul, 25
Boghossian, Paul, 54, 147-149, 150
Bonjour, Laurence, vi, 28, 29, 32, 36, 37, 38, 40, 42, 44, 45, 71, 79, 98, 99, 102,103,105,106, 119,121,122,123,124,125,126,127,154
    on rational insight, 28f
    on our knowledge of synthetic a priori truths, 36f
    on induction, 119-123.

brains in vats,
Putnam's problem, 108
Conclusions about, 135f.
Braude, Steve, vii
Brueckner, Anthony, 109, 154
Burge, Tyler, 91, 92, 93, 110, 154
Butler, Joseph, 19, 154, 155

Carnap, Rudolf, v, vii,
5,10,24,31,42,43,53,57,58,59,60,62,63,64,65,66,68,79,80,107, 127,
129,138, 139, 142,143,150,151,154,155,157,159
on analysis as reconstruction, 5, 58ff
on A-postulates, 59
on concept of analyticity, 57-60
on semantical rules, 57f

Cartwright, Richard 153-155.138, 154
Chisholm, RM., iii, vi, 19, 25, 36, 37, 38, 40, 42, 43, 44,66,71,74,77,90,93,98,101
134, 138,139,141, 156,157
on CP principle, 134, 156
on the idea of a self as a particular, 75
Church, Alonzo, 26, 79, 80, 147, 155
closure paradox, 15
Cohen, Stuart, 4,155,157, 159
coherentism, 102
colors, incompatibility of 37-40
proof that determinate colors are incompatible, 39.
concepts
as understood in this book, 85
relation to the world, 87
the word used loosely and equivocally, 81
conceptualism, 43
conditioning, rule of in probability, 131
confirmationism, 107
Crane, Tim, 98, 101

Davidson, Donald, 10
default entitlement, of logic and methodology 148
definitions
proper, rules for 149
stipulative, 60, 149
technical, form of, 58
DeRose, Keith, 4, 15, 155
Descartes, 3, 26
distributive singular terms (or DSTs), 83
Devitt, Michael, 49
Dretske, Fred, 8, 9, 15, 104, 155
Dummett, Michael, 81
Earman, John, 127, 135
Einstein, Albert, 30, 127
epistemic relativism, 4-6
epistemic justification,
internalist conception of, 102

Euclidian geometry, 30
Ewing, A.C., 44
experimental inference
and Hume, 105
and inference to best explanation, 128ff
and Bayes' theorem: *see* Bayes' theorem.
explication; *see also* analysis
as reconstructive, 59
as stipulative and complete or partial, 62, 146

Feigl, Herbert, v, vii, 154, 155, 160
Feldman, Richard, 4,10,102, 155
Field, Hartry, 54, 147-149.
Fine, Kit, 41
Fitch, Gregory, vii, 66,156
Fogelin, Robert, vii, 2 3,4,11,100,102, 143,156,160
Frege, 47, 48, 57, 61, 65, 66, 67, 68, 71, 72, 73, 78, 79, 80, 81, 83, 87, 93, 94
as interpreted by Carnap and Church, 79
on analytic-synthetic distinction, 47

Gauker, Christopher, 56
Geach, Peter, 80, 92
Gendler, Tamar, 101
Gettier, ii, 3, 11, 16, 17, 18, 19, 21, 156
Ginet, Carl, 18
Gödel, 49, 147
criticism of Russell, 35
Goodman, Nelson, 83, 123, 124, 125
and grue, 122
Gribben, John, 104
Grundstra, Bernard 123
Gupta, Anil, v, 100

Hawthorne, John, 13, 15, 101, 155, 156, 157, 158, 160
Hintikka, Jaako, 27, 29
holism
and Quine, 104
horseshoe symbol, 56
Howson, Colin, 132, 136
Huemer, Michael, 97, 98, 100
Hume, David, 22, 95, 96, 97, 100, 105, 106, 107, 109, 110, 114, 119, 121, 122, 134, 139, 140
hypotheses
"bent" 124
BIV 136
skeptical,136
hypothetico-deductive method, 127

indiscernibility
as identity condition for determinate colors, 61
individualism, 90

induction, 120
and laws, 123
Lycan on, 126
and Russell, 120
BonJour on, 122
inference
causal or experimental, 119
to the best explanation, 106, 124, 126
criticized, 126
intuitions
in Kantian sense, 79
as problematic 28-31.

Jackson, Frank, 23
Johnson, Mark, 98, 101

Kahneman, D., 126
Kalish, vii, 26
Kant, Immanuel, 22, 23, 24, 27, 35, 40, 4348, 50, 68, 71, 79, 80, 88
definition of analytic truth, 48
Kaplan, David, 23, 65, 66, 67, 68, 69, 79, 90, 107
Kemp Smith, Norman, 119
King, Jeffrey, 66, 68
Kneale, Martha, 26
Kneale, W.E.,
knowledge, 3, 4
and information, 12
and standards of evidence, 4
a priori knowledge, 22–25
a priori vs. a posteriori, 22
and the world: problems concerning, 104
as acquaintance, 2
conclusive, 3
empirical, foundation for, 99
intuitive knowledge, 27–31
uninferred, 100
senses of the word, 1
strong or perfect, 3, 20
weak sense, 3
Kornblith, Hilary, 2
Kripke, Saul, ii, v, vii, 24, 25, 42, 43, 63, 64, 66,156, 158
and necessities known a posteriori,25
and contingencies known a priori, 24f
arguments concerning the origin and composition of a thing, 42f, 63f

Laurence, Stephen, 81
laws of logic, 26ff, 32-36, 51
and rationalism, 26ff
justificaton of 147
LaPorte. Joe, vii, 12, 158.
Lehrer, Keith, 102
Leibniz, G. von, 74
Lewis, David, ii, 4, 5, 6-7, 6, 7, 8, 9,10,11,12,18, 20, 21, 55, 57, 59, 67, 71

71,76,80,102,115,158
and definition of knowing, 6
and epistemic contextualism, 6-11
and evidence for knowing, 6
linguisticism, 43
Liddell and Scott, 82
Lipton, Peter, 126
Locke, John, 73, 74, 96, 97, 134
Logic,
applied to natural languages, 53
logical truth, 47, 48, 49, 50, 52, 54, 61.
as formal notion, 51.
see also laws of logic.
logical truths, generality of, 52
logical empiricism, v, 22, ch. 3 *passim*

Lewis, David, 55, 57, 59, 67
Ludlow, 5, 158
Lycan, William, 113, 120, 126, 127, 128, 135

Macfarlane, John, 4, 14, 158
Margolis, Eric, 81
Martin, Richard M. 110
Mates, Benson, 74, 158
Meaning Postulates or A-postulates
as partial stipulations, 59, 146
Malcolm, Norman, 3
Maxwell, Grover, vii
McGinn, Colin, 41, 42
meaningful words
Hume's criterion for 106
memory
as source of empirical knowledge, 117
skeptical view of, 117
metaphysical realism, 113
Mill, J.S., 29, 30
modus ponens
apparent counter-example to, 54
validity of, 54
what is it, 56
Montague, Richard vii, 26
Moore, George Elmer, 3, 78

natural deduction, 26
Nicod, Jean, 25, 26
Nisbett, Richard E., 98

observation
problematic character of, 95
observational evidence assessed, 97
solution of Hume's problem about, 100

partial analysis
importance of, 59
particulars
as "bundles" or bare objects, 74
Peacock, Christopher, vi
Perry, John, 66
Phillips, Lawrence, 124, 132, 136
Plantinga, Alvin, vi
Plato, 1,3
postulate
compared with semantical rule, 58
predication, 76
Principia Mathematica, 25
Prior, A.N. 152
probabilities
categorical and conditional, 129
prior and posterior, 131
needed to apply Bayes' theorem
how to obtain, 135
probability
evidential and physical, 128
calculus for
axioms of 157
interpretations of, 128
probability statements
basic
how justified, 133
properties
advantages of F-theories, 78
as understood by van Inwagen, 79
problems with A- and T-theories of, 73
theories of, 72
propositional attitudes
conceptualist view of, 88
Propositional clause
oratio recta and oratio obliqua, 88
and propositions, 90
propositions
"what is said" 153
and Frege, Carnap, Kaplan, 65
as sets or functions, 67
as classical conception of, 65
Putnam, Hilary, vi, 80, 107, 108, 109, 110, 111, 112, 113, 114

quantum mechanics,
special logic allegedly needed for 149
Quine, W. V. O. ii, v, 24, 25, 42, 47, 48, 49, 50, 51, 52,53,54,57,58,
60,61,104,144,154,155, 154,159
Early and late views on analyticity, 49
Quinton, Anthony, 98
quotation
direct and indirect, 88

Ramsey, Frank, 133
RATIONALISM, epistemological, ch. 2 *passim*
vs. empiricism, 22
reasons for believing something 103
reference, Putnam on 113.
Russell, Bertrand, 1, 25, 26, 30, 34, 154
Rosenberg, Jay, 5
Rosenkrantz, Roger vii, 124, 125, 132, 135, 136
Russell, Bertrand, 71, 73, 74, 97, 98, 118, 120, 121, 122, 123, 124, 138

Salmon, Nathan, 64
Schaffer, Jonathan, 9,15,160
schematic formula
instances prescribed, 53
semantical rules
as understood by Carnap, 58
as rules of interpretation, 57
Shaffer, Jonathan, 78
Steup, Mattias, 71
Schiffer, Steven, 13
Sellars, vii, 82, 83, 86, 90, 101
Selves, nature of 101
semantic externalism
and names versus predicates, 112
and Putnam, 107
criticized, 110
Sense data 97
Sense-impression, concept of 101
Sicha, Jeff, vii, 85, 90
skeptical problem,
as raised by Putnam 14
sketch of solution, 138
skepticism, 2, 3, 7, 95, 97, 102, 105, 137
Skyrms, vii , 128, 132, 135, 136
Smyth, Herbert W., 88
Soames, Scott, 23, 65, 153
Sosa, Ernest, 29
Stanley, Jason. 13
Steup, Mattias, 16, 71, 154, 155, 156, 157, 160
Stipulation, defended 149
Suppes, Patrick, 33, 41, 58, 152

Talbot, William, 136
testimony
Locke on, 96
thought
conceptualist account and analyticity, 68
Trehub, Arnold, 113
Turing, Alan, 107, 111
Tversky, A., 126
type-token distinction, 82

Unger, 2, 3, 4, 160
Urbach, Peter, 136

vagueness, 81
validity, 51, 54, 61, 62
van Fraassen, Bas, v, vii, 57.124,126, 127,130, 141, 161
van Inwagen, Peter, 71, 79
verificationism, 107
von Wright, G. H. 149

Wetzel, Linda, 83, 84, 85
Wilce, Alexander 149
Williamson, Timothy, 2, 11, 57, 60, 98

yellow and green all over, 61

Zalta, Edward 149